Randomization as an Incentive Device: Evidence from Public Procurement of Immigrant Integration Services*

Hanna Pesola† Matti Sarvimäki‡ Tuomo Virkola§

March 20, 2025

Abstract

We examine the impact of procuring immigrant integration services from a private provider under a contract where a randomized research design serves as an incentive device. Immigrant job seekers were randomly assigned to either a private fund or public employment services, with the private provider's compensation tied to differences in average unemployment benefits and taxes between the two groups. We find that the private fund outperforms the public alternative, increasing earnings by 15%, improving job quality, and reducing the net burden on public finances by 12% over the three-year contract period. These positive effects extend to non-contracted outcomes and persist beyond the period during which the private provider's incentives were in place. The effects are particularly pronounced for high-skilled participants. Our findings suggest that procurement contracts that credibly align the incentives of providers with public sector goals can significantly improve service quality.

*This project was registered at the AEA RCT Registry (AEARCTR-0012519).

†VATT Institute for Economic Research and Helsinki GSE; Email: hanna.pesola@vatt.fi.
‡Aalto University School of Business, VATT and Helsinki GSE; Email: matti.sarvimaki@aalto.fi.
§VATT Institute for Economic Research and Helsinki GSE; Email: tuomo.virkola@vatt.fi.

1 Introduction

A fundamental question in social sciences concerns the appropriate division of labor between the public and private sectors: which tasks should be managed by the government, and which are best left to the market? In practice, this distinction is often blurred as governments procure services from private providers. Outsourcing is typically motivated by the expectation that private firms have strong incentives to maximize profits, and thus deliver services more efficiently than the public sector. However, when service quality is not verifiable, procurement contracts may incentivize cost-cutting at the expense of quality (Hart et al. 1997).

This paper shows that incorporating a randomized research design into a public procurement contract can act as an incentive device that pushes a private provider to deliver better service than the public sector. We reach this conclusion in the context of procuring integration training for immigrants. In 2016, Finland's Ministry of Economic Affairs and Employment and a private investment fund, Epiqus, launched a joint initiative—the Integration SIB—where the fund took responsibility for arranging integration services for some immigrants. A central feature of the contract was an incentive structure built on a rigorous research design. Specifically, eligible immigrants were randomly assigned to either the private fund or the public employment services. The fund's compensation was partially tied to the difference between these treatment and control groups in average unemployment benefits and income taxes over a three-year follow-up period.

We find that the private fund outperformed the public employment services as measured by immigrants' labor market outcomes. During the first three years after randomization, cumulative earnings in the treatment group were, on average, 4,500 euros or 15 percent higher than in the control group. This effect on earnings is partly driven by the extensive margin, with employment among program participants consistently over five percentage points higher than in the control group. Furthermore, the private fund was more successful in placing high-skilled immigrants in professional and higher-paying jobs and into larger and more productive firms than the public alternative. Consequently, the effects are substantially larger for more educated immigrants. We also find a larger effect for younger participants, while we don't detect treatment effect heterogeneity by gender or years-since-migration.

As immigrants in the treatment group performed better in the labor market, their tax contributions increased and the use of social benefits decreased. On average, the treatment group created a 2,700 euros or 12 percent lower cumulative net burden on public

finances over the three-year follow-up period than the control group. In total, 2,636 immigrants were randomized into the Integration SIB, 7.4 million euros were paid to the private fund, and approximately 5.5 million euros were saved in active labor market policy expenditures through public employment services. As a result, the net fiscal impact over the duration of the contract was approximately 5 million euros.

We argue that the positive effects of the Integration SIB are likely due to its contract structure, which created incentives for improving service quality. It was implemented as a social outcomes contract, specifically as a "social impact bond" (SIB), in which a private investor covers the up-front costs of delivering a service and is later compensated based on its estimated impact. At the time of its launch, it was the second-largest SIB ever implemented (OECD 2016). As such contracts have gained popularity, concerns have been raised on private partners' ability to influence impact metrics, for example, by "cream-skimming" easier-to-serve clients or using outcome measures that the providers can manipulate (Fraser et al. 2018). In the Integration SIB, these concerns were addressed by using (i) a randomized research design to limit cream-skimming and (ii) cumulative and register-based outcome measures to make manipulation unprofitable. This structure created dramatically different incentives in comparison to the business-as-usual model. In the Integration SIB, compensation was linked to its contribution on public finances, while traditional public employment services lacked performance-related incentives.

As with any contract that ties compensation to specific outcomes, the Integration SIB carried the risk that the provider might neglect other important dimensions of quality (Holmström and Milgrom 1991). In particular, the private fund had strong incentives to maximize short-term outcomes, potentially at the expense of investments that could yield higher returns in the long run. Moreover, because compensation was linked only to income taxes and unemployment benefits, the fund could have increased its financial returns by shifting participants from unemployment benefits to other forms of social benefits rather than improving their labor market prospects. We find no evidence of such unintended consequences. We track all participants for four years after randomization and the first cohort for six years, and find that the positive effects on earnings and net fiscal impact persist roughly constant over this period. Furthermore, half of the net fiscal gains stem from reductions in non-contracted transfers, showing that the program's benefits extend beyond its direct incentive structure.

Beyond showing that incorporating randomization into a contract can improve service quality, the Integration SIB also provides insights into how immigrant integration

services can be improved. To understand how the private fund's approach differed from that of the public employment office, we first draw on document analysis and in-depth interviews conducted as part of a government-commissioned evaluation.[1] The Integration SIB placed participants in short, occupation-specific language training, which aimed to provide them with the tools needed to begin working in a targeted field. Within 4 to 6 months of the program's start, they were placed in jobs or on-the-job training in these occupations. The private fund played an active role in employer recruitment, job matching, onboarding, and, in some cases, conflict resolution between employers and employees. In contrast, public employment services provided integration support primarily to recently arrived immigrants, offering extended in-class language and general training before transitioning participants to standard employment services. At the time of randomization, most participants in our sample had already completed standard integration training or were outside the three-year integration plan window. Thus, for the majority, the relevant counterfactual to the Integration SIB was standard public employment services.

Unfortunately, we do not observe in detail what services were provided to participants by the private fund. However, we do have detailed information on the services offered by public employment services (PES). These data allow us to characterize the counterfactual services and thus provide hints about the potential mechanisms behind the program's effects. We find that high-skilled immigrant job seekers are less likely to receive job search assistance in the form of vacancy referrals from PES than low-skilled immigrants, and when they do, the referrals are often to low-skilled jobs for which they are overqualified. They also participate less frequently in general training or subsidized employment programs compared to low-skilled immigrant job seekers but spend more time in in-class language training. Since high-skilled immigrants experience the largest gains in earnings and job quality from the Integration SIB, this suggests that the private fund was particularly effective in improving service quality for this group. A likely explanation lies in the differences in economic incentives: given their higher earnings potential, it was more valuable for the private fund to invest resources in high-skilled job seekers and to match them with jobs that better aligned with their skills.

Finally, we examine the extent to which our results could follow from the treatment group getting jobs that would have otherwise been filled with members of the control group. We leverage variation in program roll-out across labor markets and find no indica-

[1] This evaluation consisted of a quantitative part—an early version of this paper—and a qualitative part, conducted at the Owal Group by Risto Karinen, Nella Koivula, Tiina Koivula, and Olli Oosi. An English translation of the report is available as Karinen et al. (2024).

tion for the presence of such displacement effects. We argue that important displacement effects are unlikely, given the limited size of the experiment. Even in the most affected region, Uusimaa, the experimental group responded to 0.3 percent of the working-age population and 4.7 percent of working-age immigrants. Furthermore, our results suggest that the positive effects were partially created by helping immigrants enter parts of the labor market that they would otherwise not have had access to. Thus, displacement effects are unlikely to bias our estimates substantially.

Our primary contribution is to the empirical literature on outsourcing public services and contract design. The existing literature on outsourcing highlights the importance of context, both in the characteristics of the contracted services and in the broader institutional environment in which they operate (see Andersson et al. (2019) and Fabre and Straub (2023) for reviews). Prior work suggests that outsourcing can lead to efficiency gains when service quality is easily verifiable, such as in garbage collection (e.g., Szymanski and Wilkins 1993; Szymanski 1996; Gradus et al. 2018). However, the evidence is mixed for services such as health care, education and prisons, where contracting on quality is more challenging. Many studies find reductions in service quality (e.g., Chan et al. 2023; Knutsson and Tyrefors 2023), while some find no effects (e.g., Duggan et al. 2018), and others report improvements in quality (e.g., Bergman et al. 2016; Romero et al. 2020). Closest to our study, evidence on the private provision of active labor market programs suggests that outsourcing these services does not consistently result in cost savings or improve labor market outcomes for unemployed job seekers (Bennmarker et al. 2013, Krug and Stephan 2016, Behaghel et al. 2014, Rehwald et al. 2017; see Crépon 2018 for a comprehensive review).

We appear to be the first to examine a public procurement contract in which compensation is determined using a randomized research design. We also provide the first analysis of outsourcing entire employment services rather than purchasing specific training courses from private providers. We recognize that linking compensation to a provider's performance relative to a peer group is an old idea going back to at least the widely used yardstick competition approach (Shleifer 1985). A key difference in our approach is using a randomly selected control group, which provides a more reliable benchmark than those used in conventional yardstick competition applications. This improvement is crucial because a more accurate performance signal reduces noise and thus makes it easier to assess the provider's true contribution. Consequently, it strengthens the provider's incentives to put more effort into improving service quality (Holmström 1979; Grossman and Hart

1983).

An additional advantage of embedding a randomized research design into the Integration SIB is that it allows us to inform the literature on policies to improve the labor market integration of immigrants (Åslund and Johansson 2011, Joona and Nekby 2012, Sarvimäki and Hämäläinen 2016, Battisti et al. 2019, Lochmann et al. 2019, Dahlberg et al. 2024, Heller and Mumma 2023, Foged and van der Werf 2022, Arendt et al. 2022, Foged et al. 2022, 2024; Pesola and Sarvimäki 2024). This body of work shows that integration programs tend to have large positive effects on participants' and their children's outcomes. However, current evidence is limited to programs focused on language training and targeted to recently arrived immigrants, typically refugees.

By contrast, we examine integration policies focusing on job placements for immigrants who had, on average, already lived in the host country for seven years. Importantly, unlike the programs studied in the earlier work, part of the services in the Integration SIB were tailored explicitly for highly educated immigrants, and our findings show that the training was particularly effective for them. In addition, the availability of a large-scale randomized research design sets our paper apart from most of the earlier work.

The rest of the paper is organized as follows. Section 2 describes the Integration SIB program, and the next two sections discuss our data and empirical approach. Section 4 reports our results and section 5 discusses potential mechanisms behind them. We end with some concluding remarks.

2 Institutional Background and the Integration SIB

This section describes the intervention and our experimental design. We begin by providing some institutional background on the services provided by the PES to the control group. We then discuss the features of and the assignment to the Integration SIB program.

2.1 Counterfactual: Public Employment Services for Immigrants

The control group in the Integration-SIB programs participates in business-as-usual services provided by the PES. The PES offers job search assistance and active labor market programs (ALMP) to all unemployed job seekers in Finland. In addition, recently arrived immigrants who have been residents for less than three years are entitled to integration

services such as language training and other courses aimed at improving their access to the labor market. Following integration training, the PES may encourage and subsidize immigrants without secondary education to enroll in degree programs (typically vocational education) or to participate in other services offered to other unemployed job seekers. These services can include job search assistance, in-class and on-the-job training, unpaid trainee positions and wage subsidies for employment. Job seekers are assigned to training courses by the PES and, if assigned, have to participate in order to be eligible for unemployment benefits. Services and training provided by the PES are free of charge for job seekers.

As a key part of the program that we study is a change in financial incentives, it is useful to understand what they are in the control group. The PES procures training from both public educational institutions and private providers. In principle, the procurement process forces providers to compete against other providers both in price and quality. However, after winning a bid, the financial incentives are typically only weakly performance-related. First, training providers are typically compensated based on the number of customer training days. While this does incentives providers to decrease drop-out rates, it is unclear how it affects service quality. Second, neither the PES or the training providers have access to information on their customers' earnings or labor market trajectories following service participation. This limits their ability learn and improve on the services in a way that would improve participants labor market integration.

2.2 Intervention: Integration SIB

The Integration SIB -program has two key differences with the business-as-usual services provided to immigrants who are unemployed job seekers.

First, in the Integration SIB treatment, a private provider takes on the responsibility of arranging services to unemployed job seekers and is compensated for cost savings that it is able to generate relative to job seekers in the business-as-usual case. The key innovation in the program is that, as participation in the program is randomized, the cost savings can be credibly estimated from actual differences between individuals who are assigned to participate in the program and those that were not. In practice, these cost savings are calculated as the differences in net transfers (income taxes paid - unemployment benefits paid) to the job seekers in the treatment group vs the control group over a three year period, and the service provider gets 50 percent of this cost saving. Importantly, the randomization process guarantees that the service provider cannot game the system or

do cream skimming by selecting participants who it expects to be most likely to find a job irrespective of the service quality. Thus the program creates strong economic incentives for service providers to help the job seekers find jobs at least during the three-year period.

Second, the program also includes qualitative differences in the services provided to the job seekers. The participants in the Integration SIB initially take part in occupation-specific language-training after which the aim is to directly match them to jobs with employers that the service provider has been able to make contracts with. Typically these jobs were marketed to be in entry-level occupations that were experiencing labor shortages and which had moderate language requirements. Thus, the main qualitative difference is that job seekers in the integration SIB are expected to move to on-the-job training and employment relatively fast. This is in contrast to the business-as-usual, where unemployed immigrants are expected to spend more time in in-class language training and more general training rather than on-the-job training.

Participation in the integration SIB was voluntary and publicly marketed to immigrants who registered as unemployed job seekers. These job seekers were likely to learn about the program in a meeting with a PES caseworker. The PES also handed out brochures to the target group and organized promotional events to attract individuals to apply. There was a website which described the main aspects of the program, the eligibility criteria and instructions on how to apply.

2.3 Randomization

Starting in Fall 2017, unemployed immigrants who want to participate in the integration SIB first approach the service provider Epiqus indicating that they would be interested in participating. The service provider interviews the applicants and selects suitable candidates. The service provider then sends a list of candidates to the Public Employment Service (PES) who verifies that all individuals on the list fulfill the program's eligibility criteria. The PES then places eligible individuals on a list and reorganizes the list to a random order using the website random.org. Then, the PES assigns the first 70 to 80 percent of individuals on the list to the treatment group and the rest to the control group. Among applicants without tertiary education, 70% were randomized into treatment and 30% into the control group, while among applicants with tertiary education 80% were randomized into the treatment group and 20% into the control group. The randomization takes place regionally at the beginning of each week and is done separately for individuals who are subject to integration plans (have been in Finland for less than three years) and those who

are not (have been in Finland more than 3 years). The PES then sends the list of applicants who were accepted to the program back to the service provider who then invites the accepted individuals to the program.

Individuals can apply to the program and are included in the randomization only once, and hence the result of the randomization is an absorbing state. In practice, after the randomization, the PES marks the randomization outcome in their system indicating whether the individual was randomly selected to the treatment or the control group. This information is used in the following rounds of randomization to check the applicants eligibility. Individuals can only apply once, and if they are assigned to the control group, they are not eligible to re-apply on a later date.

3 Data and Research Design

We use administrative data from the Public Employment Service (PES) on all eligible applicants to the Integration SIB-program between 2017 and 2019. In our sample period, eligible applicants consist of 17 to 63 year-old immigrants, who do not have Finnish citizenship at the time of program assignment, and who have registered as unemployed job seekers at the PES. Applicants are required to be able to read and write in their own language. In total, between 2017 and 2019 there were 3662 eligible participants of which 2634 where randomized into the treatment group and 1026 into the control group.

We link the data on program participants to Statistics Finland's population wide data from several administrative registers. The data contain information on everyone who is a resident in Finland and their demographic characteristics (birth dates, gender, Finnish educational degrees, country of birth, native language, and immigration date). These data also contain details of employment spells, employers, occupations for workers and industries for employers. In addition, we use comprehensive information on earnings, taxes, social security contributions paid and benefits received from the Social Insurance Institution and the Tax Administration. Finally, we use further data from the PES register on job seekers, including data on unemployment spells, participation in active labor market programs and certified language skills for immigrants.

Earnings Our primary pre-registered outcome is annual labor earnings which is based on tax data. We have data until 2023 which allows us to measure earnings for up to 4 years after program assignment for all participant cohorts and up to 6 years for the first cohort.

The benefits of annual earnings is that it captures both extensive and intensive margin effects and provides a useful summary measure of labor market status of migrants for whom both margins are likely to be important. We concentrate on the earnings responses in the first three years after program assignment, which corresponds to the period for which incentives for the private fund are in place. However, we follow earnings also as a function of time from assginment to provide more evidence on the potential mechanism and the potential short- and long-run trade-offs..

Net Transfers To understand whether the fund was able to improve on the outcomes of the public employment services as specified in the procurement contract, we use tax and transfer data at the individual participant level. We first decompose net transfers to tax contributions and unemployment benefits which were jointly determining the compensation for the fund. However, we also study other, non-contracted transfers to get a more comprehensive estimate of the direct fiscal effects of the program and to understand potential spillovers to other benefit programs.

Employment We use information on work contracts to construct measures of employment. We count both work days during the year and monthly indicators of being employed in any firm.

Job-quality We use information on occupations to job quality. These data are available for each end-of-the-year job contract. Based on ISCO-classification, we first group jobs into professional (1-3), clerical and service (4-5) and manual (6-9) jobs to separate jobs that clearly have different skill requirements and earnings prospects. Second, we use information on the population level to measure the quality of jobs with more granularity. In particular, we construct expected earnings and skill-requirements using information on the earnings and higher education status by workers in each 4-digit occupation in the prior 5 years before the program.

Training and subsidies To measure participation in PES provided programs, we use individual level information on course and training participation. We classify these programs to four categories: immigrant training, general training, employment subsidies (including unpaid work trials and wage subsidies) and PES provided education subsidies. The first of these is targeted to immigrants and mostly to immigrants in the first

three years after arrival. All others are targeted to all job seekers, but conditional on the discretion of the PES and the job seeker.

Vacancy Referrals To measure the job search assistance the participants receive from the PES we use information on vacancy referrals. We count the number of referrals each participant receives to quantify their importance. These referrals also contain information on the occupation and firm in which the jobs is, which allows us to measure what types of jobs the PES recommends to job seekers and to what extent these match the skills of the job seekers.

3.1 Research Design

For our main outcomes, we estimate the effects of the treatment based on

$$Y_{it} = \alpha + \beta_s \text{Treatment}_i + \theta_{j(i)} + \lambda_t + \gamma X_i + \varepsilon_{it} \tag{1}$$

where Y_{it} is the outcome of interest observed at year t, $\theta_{j(i)}$ is a fixed-effect for randomization event in which individual i's randomization takes place, λ_t is a year fixed-effect and X_i is a set of individual level controls. In the baseline analysis, X_i includes age, gender and a dummy for the individual being subject to integration training. In addition, to characterize the differences in trajectories over time and to understand the potential mechanisms, we also present exploratory analysis using event-studies where we interact the treatment status with time since program start.

In all of our analysis, we focus on intention to treat estimates as we don't have perfect information on which of the randomly invited participants started in the Integration SIB treatment or received services from the private fund. However, we do have information on participants who participated in a training that fulfilled the 3-month training criteria and allows the fund to claim the 1,500 euro payment from the pubic sector. In total, 70 percent of participants fulfilled this criteria, suggesting a high take-up. However, some participants might have received services for a shorter duration or services that did not fulfill the training criteria. Thus, this information likely underestimates the true take-up and consequently biases instrumental variables approaches.

3.2 Descriptive Statistics and Balance Tests

Table 1 shows pre-treatment descriptives of the participants and balance between control and treatment groups. The average participant is 39 years old, has been in the country for almost 7 years and registered as an unemployed job seeker for more than half a year. Annual earnings are 4,000 euros on average, and both treatment and control groups receive about 10,000 euros in social benefits, mostly from unemployment benefits. This results in negative net transfers of more than 8,000 euros to the public sector. Further, there are no important differences across the treatment and control groups, as expected due to random assignment.

4 Main results

In this section, we present our main results. We start by characterizing the effects on immigrants' earnings and employment which were the main targets of the program. Second, we characterize to what extent the fund was able to improve on the metrics agreed in the procurement contracts and whetter or not there were spillovers to other government welfare programs. Third, we discuss issues related to potential displacement effects that could plausibly affect the interpretation of the estimates. Finally, we discuss measures of job quality and heterogeneity in the estimated effects by participant characteristics.

4.1 Labor market integration

Figure 3 shows the effects of the Integration SIB-program on earnings. Figure 3a shows trends around program assignment separately for the control and treatment groups. It shows that, on average, participants had annual earnings of only around 4000 euros before being assigned to the program. This illustrates the poor labor market performance of the participants who are primarily unemployed job seekers. Following the assignment, however, earnings increase significantly in both the treatment and control groups, but they increase persistently more in the treatment group than in the control group. Figure 3b shows the differences between the treatment and control groups in an event-study design, controlling for cluster and year fixed effects. The earnings effects are around 1,500 euros per year in the years after being assigned to the program.

Taking these together, Figure 4 shows the results on our pre-registered primary outcome which is the total earnings in the first three years of the program. This is the time

during which the service provider was responsible for providing the services to the participants and had financial incentives to improve their labor market status. We find that the effect on earnings is an increase of 4500 euros or 15 percent, and this is clearly statistically significant. Thus, the program was successful in achieving its target of improving labor market outcomes of its participants.

To understand whether these earnings effects are driven by extensive or intensive margin, we next use information on employment contracts to follow individual employment trajectories. Figure 5 shows both the trends and an event-study design. Panel (a) shows that the participants are on a downward trajectory and less than 20 percent had any employment contract at the time of the assignment. Both treatment and control groups are more likely to become employed in the 6 months after assignment after which the improvement in employment slows down. However, the treatment group is significantly more likely to be employed. The point estimates from Panel (b) are consistently above 5 percentage points (or 15 percent) higher than in the control group. The gap narrows somewhat after the first 30 months but remains positive at three percentage points. Taken together, the results on earnings and employment imply that earnings improve more also among the employed in the treatment group than in the control group.

The positive earnings and employment effects are also very persistent. The program gave credible incentives for the private provider to support participant labor market integration up to three years, but little incentive to do so thereafter. This could have resulted in participants being employed in temporary jobs or other arrangements that might not benefit the participants in the long run. Yet, we see that both earnings and employment remain higher still in the fourth year. which is the longest window we have available for all the treatment cohorts. For the two cohorts that we observe for more years, we also see no drop in the earnings responses, suggesting that the earnings effects remain significantly persistent for multiple years.

4.2 Net transfers to the public sector

To evaluate to what extent the private service provider was able to improve the contracted outcomes, we next turn to net transfers to the public sector.

Figure 6 shows a decomposition of net transfers to the public sector during the first three years after program assignment. The first column shows the amount of taxes (and social security contributions) paid by participants and by treatment status. On average, the treatment group paid 1330 euros more than the control group in taxes as predicted

by their increased earnings. The second column shows the unemployment benefits collected. Here the differences between two groups are small, around 21 euros or less than a percentage point reduction for the treatment group.[2] Taken together, the reduction in taxes paid and small change in unemployment benefits resulted in a total of 1350 euros (15 percent) reduction in net unemployment benefits paid to the treatment group (Column 3). Thus the private providers were able to create savings to public finances in the contracted measures.

A potentially critical issue with the contracted measures is that as they do not contain all social benefits or costs to the public sector. For example, the private provider could prefer participants with low employment prospects to move outside the labor force and to take-up other social benefits such as social assistance. To evaluate these potential spillovers, the columns 4 and 5 show the differences in non-contracted social benefits and total net transfers that include also these additional social benefits. Column 4 shows that there was a 1321 euro (9 percent) reduction in non-contracted social benefits. Thus, the program actually created positive spillovers and more savings to public sector than what was contracted. The total reduction in net transfers was thus 2671 euros (12 percent), almost double what the private provider was compensated for.

The above results suggest that the incentives were set both credibly and in the right direction. However, the fact that the contract only compensated for part of the savings means that the incentives could have been even stronger if all transfers were included. However, with this research design we can only speculate whether the effects on labor market integration would have been even larger if these incentives were stronger. It is also possible that the service provider could also learn about the effects on other transfers ex post as these were not directly observable to the service provider.

4.3 Displacement

Previous evidence suggests that large scale active labor market programs can have displacement effects. Now, the Integration SIB -program was relatively small in size relative to the labor force which makes it unlikely that the effects that we estimate would be purely driven by displacement. However, to aim to gauge the role and potential magnitude of such displacement effects, we leverage the differential roll-out of the program across labor markets (see Table A1 for descriptives). The program initially started in the Uusimaa

[2]The small effect on unemployment benefits can be explained by reductions in being outside the labor force but also due to higher levels of income-dependent unemployment benefits after spells of employment.

region which covers the Helsinki metropolitan area and surrounding municipalities and then rolled out to other regions. The great majority of participants (82 percent) were participating in Uusimaa and thus that area had the highest treatment intensity relative to the size of the labor force. Overall, 3.6 participants per thousand job seekers were part of the . In comparison, the treatment intensity varied from 1.1 participants per thousand to no participation in other regions in the rest of Finland.

Table 2 Panel A shows our estimates on treatment effects by the scale of the program in the labor market. First Column (1) shows estimates using variation in the intensity linearly interacted with the treatment and Columns (2) shows estimates comparing Uusimaa to the rest of Finland. The point estimate for both specifications is negative and thus inconsistent with displacement effects being the driver of our main results. If there were significant displacement effects, we would expect the treatment effects to increase as a function of its size. However, the point estimate is not significant. Also, we find that the effects are positive and significant in both the Uusimaa region and the rest of Finland if we estimate these completely separately (Table 2 Columns 3 and 4). In the Appendix Figure A1, we also show earnings and employment trends in both Uusimaa and the Rest-of-Finland, both of which show similar patters by regions.

4.4 Quality of Jobs

The program marketed job opportunities in low-skilled occupations with labor shortages. A natural concern that this may raise, is that these are jobs from which occupational mobility and career progression could be limited. As the participants spend less time in general language training and general education, it is at least plausible that they might risk being stuck in entry-level jobs in the long run. To understand whether this is the case, we first use information on occupations in the registry data. We classify jobs to three broad occupational groups "manual", "service" and "professional" based on the first digits of the ISCO classification.

Figure A2 shows evidence on the types of occupations that the immigrants in our sample hold. First, as predicted, most work in manual low skilled jobs and service sector jobs after the program starts. After three years, 14 and 24 percent of participants are working in manual or service occupations respectively. Somewhat surprisingly, however, there is no important difference in holding these low skilled jobs between the treatment and control groups in the longer run. In contrast, the treatment group is significantly more likely to hold higher level professional occupations. Thus, it does not appear that the jobs that

participants were matched were worse than in the counterfactual or that that the participants would be particularly stuck in low-skilled occupations. Rather, it appears that these matches provided opportunities to access occupations with higher skill-requirements.

More formally, we study the effects of the program on the quality of occupations during the first three years in Table 3 Panel A. Columns (1) to (3) show that indeed the increase in employment is mostly driven by increases in professional jobs rather than service or manual jobs. The treatment group is 2.3 percentage points more likely to hold a professional job relative to the control group. This is a 31 percent increase over the sample average. The effects on service and manual jobs are between 0.7 and 0.5 percentage points and smaller relative to the baseline. We can also characterize the quality of occupations more granularity by summarizing them with average earnings for Finnish workers in those occupations (Columns 4) or the share of Finnish workers in each occupation that have a higher education degree (Column 5). Both of these measures point consistently in the same direction: the treatment group works in better jobs with higher skill requirements than the control group. The program increased expected earnings by 1230 euros (4.2 percent) and the share of colleagues with higher education increased by 2.8 percentage points (17 percent).

To characterize the jobs and workplaces in complementary dimensions, we look both at co-workers and firm level productivity measures. These are consistent with the above results. First, we find that co-workers earn around 1500 euros or 5 percentage points more (Columns 1 and 2 in A2 Panel A). Second, we use an AKM decomposition to characterize firm-level wage premiums and coworker skills. Overall this decomposition suggest that the participants work in firms that pay 7 percentage point lower wages to workers and their coworkers are paid 22 percent less than the average Finnish worker. However, treated workers are not working in significantly higher pay-premium firms than the control group (Columns (3) and (4) in A2 Panel A). Third, using firm registry information, we can construct a measure of firm productivity using sales and the number of workers in the firm. We find that the treated workers are working in more productive firms. As sales per worker may vary significantly by industry, we also provide estimates that control for the industry.[3] Our more conservative estimates suggest that the treatment group works in firms with 5.5 percent higher productivity.

[3]In Appendix Figure A3], we show that the treatment group is more likely to work in ICT which has more professional workers as well as in manufacturing, retail and accommodation and food services. In contrast, there is a drop in the share of jobs in administrative and support services which includes cleaners and other elementary occupations.

Overall, our results point clearly in the direction that the program participants are finding better jobs than those in the control group. This is despite the increase in employment. Thus, the increase in employment is not coming at the expense of job quality.

4.5 Heterogeneity by Skill and Potential for Targeting

Next, we turn to analyzing how the program impacts vary by participant characteristics. We use this as an opportunity to learn about the potential mechanism. For policy making, it would also be important to identify groups that would be most likely to benefit from new programs and target the programs to those population groups. To make some progress, we next look at how the effects on our main outcomes vary by time since arrival, gender, age and education.

Table 4 and Figure 7 show the differences in earnings and employment trajectories across different groups. Columns (1) and (2) in Table 4 and Figures 7a and 7b show that the effects of the program on earnings and employment are remarkably similar for more recent immigrants and those who have been in the country longer, as well as for men and women. There do, however, appear to be differences in the earnings impact of the program for immigrants of different ages. Column (3) in Table 4 and Figure 7c suggest that the earnings effects are highest among young workers. The fact that younger workers benefit from the program more and the effects are more persistent, could suggest that the program was able to smooth frictions in accessing entry level labor markets for groups that have no significant prior labor market experience.

The clearest differences in the effects of the program emerge for immigrants with different levels of educational attainment. Based on the results in Column (4) of Table 4 and Figure 7d the program was especially effective for immigrants with higher education. The effect on low-skilled immigrants earnings is around 2200 euros, but over 1000 euros for high-skilled immigrants. This effect is mostly driven by the intensive margin rather than the extensive margin, as the effects on employment are not statistically significant and relatively small. This suggest that the program was succefully placing high-skilled immigrants into better jobs. Indeed, in Table 3 Panel B, we show that the effect on job quality in terms of occupations and occupational skill requirements is mostly driven by better jobs to high-skilled immigrants.[4] We explore potential mechanisms behind these heterogeneous effects in section 5.

[4]In the Appendix Tables A2 A3, we also show that high-skilled immigrants are driving the positive effects on other measures of workplace quality.

5 Mechanisms

The Integration SIB -program introduced credible financial incentives to the private fund to provide services that help participants integrate into the labor markets. In this section, we aim to make progress in understanding the mechanism through which these positive effects come about. Specifically, we aim to understand what the fund did differently than the PES does in the business as usual services.

5.1 Job Search Assistance

We showed above that participants with higher education benefited more from the program than those without. Their earnings increase more (Table 4) and they are more likely to attain professional jobs that pay more and which have higher educational requirements (Table 3 Panel B). This is not primarily driven by the extensive margin, as participant education does not predict significantly more work days. This could suggest that the business-as-usual services are not able to help skilled immigrants as well as low-skilled immigrants. Or, alternatively, that high skilled immigrants themselves can benefit more from additional services than low-skilled immigrants.

One explanation for this finding might be that, having more earnings potential, the private fund puts more effort into helping high-skilled immigrants find jobs that match their skills. As the private fund is compensated for every increase in income taxes paid by workers, it may have a stronger incentive to match high skilled workers to good jobs than the PES where caseworkers compensation is not tied to the quality of jobs their customers get. Now, unfortunately, we do not directly observe how the private fund allocated resources at the individual level to directly test this hypothesis. However, we do observe interactions between the PES and individuals in the control group, which can help shed light on the counterfactual services that immigrant job seekers receive outside the program. One of the important ways in which the PES can help individual job seekers is via vacancy referrals. These are recommendations to apply for a specific job opening that would be suitable for the job seeker's skills. Using individual data on these referrals we can quantify how much help job seekers typically receive when they are within the PES services.

Figure 8 shows the cumulative number of referrals in the 12 months after applying to the program by applicant's education. Strikingly, there emerges a significant gap in referrals by skill groups. Job seekers with a bachelor's degree receive on average only 0.21

referrals but job seekers without a bachelor's degree receive 0.49, or more than double the number of referrals. This suggests that low-skilled immigrants receive more support from the PES than the high-skilled immigrant job seekers. This is not explained by higher skilled job seekers finding jobs more quickly as the employment rate shown on the right hand panel in Figure 8 shows that the employment follows very similar patter independent of skill level.

Table 5 decomposes the referrals by type, leveraging information on the types of jobs these referrals are made. We focus on referrals during the first 6 motnsh when the gap between high-skilled and low-skilled referrals emerge and less sensitive to negative selection by those who do not find jobs in the lon run. First, in Panel A, we show that 65 percent of the referrals are made to low skilled manual and elementary jobs. This matches the realized jobs held by the control group, sugegsting that either the vacancy referrals are well targeted on average or that they direct workers job search to these types of occupations. However, when we split the vacancy referrals by job seekers skill, there is a striking similarity in the referrals. 52 percent of of vacancy referrals to high-skilled imigrants are to low-skilled manual and elementary jobs, which are clearly

Taken together, the information in the vacancy referrals for the control group suggest that high-skilled immigrants are receiving significantly less job search assistance from the PES caseworkers than are the low-skilled job seekers. When they do receive this support, it directs high-skilled workers to apply for low-skilled positions for which they are likely overqualified. Now, it could be that the recommendations to low-skill jobs is driven by language barriers which make it harder for immigrants to access jobs that match their skills. However, the increase in professional jobs among the high-skilled immigrants in the treatment group suggest that the private fund was better in assisting in the job search process for the same group.

5.2 ALMP Service Paths

In addition to referrals and support from the caseworkers, the PES can assign job seekers to active labor market programs. To understand how the Integration SIB affected the types of active labor market programs individuals were subject to, we leverage administrative data from the public employment services which cover a rich variety of different types of services job seekers participate in. Figure 9 shows results on active labor market program participation for both the treatment and control group. First, panel (a) shows the probability of participating in any ALMP by treatment status around the program assign-

ment. Participation in training is high among both control and treatment groups, but the profiles look very different. On the one hand, during the six months after assignment, the treatment group is significantly more likely to be participating in ALMP as they are participating in the intensive language training program prior to the start of job placements (Figure 9b).[5] On the other hand, the control group is around 5 percentage points more likely to be participating in business-as-usual ALMPs after that 6 month training period and this gap is persistent.

The business-as-usual ALMPs that the control group participate in during the years following program assignment consist of several different services. Panel (a) in figure 10 shows that during the first six months about 10 % of the control group participate in integration training and language training targeted at immigrants, and a similar share are employed with a wage subsidy. General training not specifically targeted at immigrants is almost as common, but both immigrant and general training programs as well as subsidized employment become less prevalent after the first year. Panel (b) in figure 10 shows that, on average, the control group participates in these services for just over two months.

The most common service after the first months is, however, education programs provided by educational institutions, i.e. education and training not procured by the PES but approved by the PES caseworker. These would typically be studies that are part of a vocational or university degree. Participation in this type of education increases steadily among the control group during the first 18 months after assignment, and after three years 10 % of the control group are still enrolled. The prevalence of enrollment in these education programs leads to, on average, five months spent in education (panel b). Thus, while the services provided to the control group include subsidized employment, the different types of in-class training are much more common, and participation rates remain at higher levels than in the treatment group.

Given that the PES appears to focus on less-educated job seekers when allocating vacancy referrals, they may also be targeting other services differently based on job seekers' skills. To study this, in panel (c) of figure 10 we show the cumulative months spent in different PES services during the first 12 months after assignment split by the education level of the job seekers in the control group. The time spent in training and other services overall is similar between job seekers with and without a bachelor's degree, i.e. the

[5]For the treatment group, we only observe an indicator on whether or not the individual is participating in subsidized training and we cannot separate whether it is provided by the private fund or some other course provider, However, the time profile clearly shows that the initial spike is driven by the private fund's training in the first 6 months as expected.

more educated job seekers appear to have access to these services to the same degree as less-educated job seekers. The more highly educated do appear to participate more in language training and other training targeted specifically to immigrants, which may be expected if their education makes them otherwise more able to enter for the Finnish labor market than their less-educated counterparts.

Taken together, the data on PES supported services and training suggest that the counterfactual services are broadly similar to those offered by the private fund. However, in the long run, many job seekers are directed towards general education programs.

5.3 Discrimination

To study whether the effects of the program could operate through reduced discrimination, we analyze heterogeneity in the program's impact along ethnic lines. In Table A4, we show the estimates from a model that interacts the treatment status by the participant's citizenship. Specifically, we estimate interactions with the 10 most common countries of origin and keep the rest of the sending countries in the comparison group. The coefficients are in descending order of the share of participants from each sending country. Overall, in Column (1) there is significant variation in the point estimates on earnings by sending country. For example, Estonian, Indian, Turkish and Iranian participants get between 4,500-7,500 euros more in total earnings during the program period relative to the more than those in rest-of-the world sending countries who receive on average 5. In contrast, Russian, Thai, Chinese and Afghan participants receive between 7,200-13,400 euros less than the rest-of-the world. Similar variation by sending countries emerge if we look at net transfers or days of employment.

Overall, looking at this variation, it would be hard to attribute this variation to pure ethnic discrimination as both negative and positive estimates are found in countries with significant ethnic heterogeneity. For example, there are caucasian Europeans (Russians, Estonians) and muslim majority (Afghan, Iranian) ethnicities represented on both negative and positive sides of the estimates. However, the estimates are also not precise enough to generally detect statistically significant differences. Yet, also the fact that, on average, the immigrant composition among colleagues remains constant, suggests that most participants are hired by firms that also typically hire immigrant workers.

6 Short vs. Long Run Trade-offs

The contract in the Integration SIB program clearly incentives the private fund place participants into jobs for the duration of the three year follow-up period. Our evidence suggest that it was indeed successful in this regard. However, an important non-contracted outcome that the government should care about is the effect on long-term labor market integration and impact on public finances.

To make progress in understanding whether there could be a significant trade-off, we begin by leveraging information on earnings outcomes for the first cohorts in the program for which we have a longer follow-up period. In Figure A4, we show that the earnings in the treatment group remain persistently higher than in the control group still five and six years after the program begun, or 2 and 3 years after the financial incentives for the fund were in place. This suggests that long-term integration is not in conflict with the contracted incentives. Indeed, it seems unlikely that a positive three-year labor market attachment would fade out immediately, at least if the positive outcome was not artificially inflated by the service provider. In our context, it is very unlikely that the providers could manipulate our main registry-based outcome measures.

Second, we might expect that the trade-off between short and long-run integration would be more important for immigrants who have just recently arrived into the country and who have not had opportunities to learn the Finnish language or to have participated in general education necessary to access local jobs. Figures A5 and A6 illustrate the differences in these counterfactuals for these two groups using data on ALMP participation. First, immigrants who have been in the country for less than three years are much more likely to have been participating in immigrant training programs immediately prior to participating in the integration SIB program. Second, the effect of assignment to the Integration SIB has a large negative effect on participating in in-class immigrant training in the 12 months after assignment among the recent immigrants but a negligible effect among immigrants who have been in the country for more than three years. We already explored heterogeneity in earnings and employment effects by immigrants who recently arrived and who have in the country longer than three years. On average, we see no significant differences across these groups. In addition, we would expect that if there is a trade-off, the labor market effects for the recent immigrants would be less persistent than for immigrants who have been in the country longer. However, we do not see evidence for this mechanism either.

Although our results suggest that there were no medium-term trade-offs between

work-first and language training participation, this does not rule out the possibility that, over the life cycle, some of the participants would have benefited more from general education. For example, without a host country degree, it might be hard to access many good jobs in the long-run. Indeed, in previous work (Pesola et al. (2024)), we show that a significant share of immigrants enroll in vocational education programs, and the results for the control group in figure 10 indicate that enrollment in education is supported by the PES. While there is no reliable data on the effectiveness of this type of adult education for this group, it is potentially important to understand to what extent the Integration SIB program affected general educational enrollment.

Figure 11 shows our estimates of the differences in general education enrollment. Among both the control and treatment groups, enrollment in education is common. In the control group, almost 25 percent are enrolled in mostly vocational education. In the treatment group this is up to 5 percentage points or 20 percent lower. Now, if education enrollment and the degrees attained lead to better employment prospects in the very long-run, this potentially creates some trade-offs that policy makers should consider before widely adopting new labor market programs. However, this still remains a conjecture, while the medium term positive effects of the new program are based on an RCT.

7 Conclusions

Governments often procure services from private providers, driven by the expectation that private actors can innovate and operate more efficiently than the public sector. However, the inherent incompleteness of contracts often leaves private providers with stronger incentives to cut costs than to preserve or improve quality. This challenge has an interesting parallel with the problem of causal inference in scientific research. At its core, quality in this context reflects the extent to which a provider's services influence outcomes valued by the government. Randomized experiments have long been regarded as the gold standard for addressing such questions in scientific research. Here, we show that the same principles can also improve the design of procurement contracts.

To the best of our knowledge, Finland's Integration SIB was the first instance where a randomization protocol was embedded within a public procurement contract. The results are promising. Immigrants assigned to the private provider experienced 15% higher earnings, improved job quality, and a 12% reduction in the net burden on public finances compared to those served by public employment services. Notably, these positive effects

extended to noncontracted outcomes and persisted beyond the contract period. High-skilled immigrants reaped the greatest benefits, with significant gains in both earnings and job quality. These findings demonstrate that well-designed incentives can drive significant improvements in the quality of employment services. Similar approaches could also be applied to other types of contracts.

Embedding randomized research designs into contracts also offers the additional benefit of generating valuable information externalities by producing highly credible experimental designs for research. In the case of the Integration SIB, our findings offer two key lessons for the design of immigrant integration programs. First, there is significant scope to improve outcomes for highly educated immigrants, a group often underserved by conventional public employment services. Second, tailored language training for immigrants who have resided in the host country for several years can deliver substantial benefits.

Finally, this paper highlights the value of close collaboration between implementing organizations and researchers. The core design of Finland's Integration SIB contract emerged from a negotiation in which one of the authors was invited as an expert on integration programs. Unexpectedly, his most influential contribution did not stem from domain-specific insights but from expertise in running randomized field experiments and leveraging administrative data in research. Creating opportunities for such serendipitous interactions between policymakers and researchers could also be valuable in other contexts.

References

Andersson, F., H. Jordahl, and J. Josephson (2019). Outsourcing Public Services: Contractibility, Cost, and Quality. *CESifo Economic Studies 65*(4), 349–372.

Arendt, J. N., C. Dustmann, and H. Ku (2022). Refugee migration and the labour market: lessons from 40 years of post-arrival policies in Denmark. *Oxford Review of Economic Policy 38*(3), 531–556.

Åslund, O. and P. Johansson (2011). Virtues of SIN: Can intensified public efforts help disadvantaged immigrants? *Evaluation Review 35*(4), 399–427.

Battisti, M., Y. Giesing, and N. Laurentsyeva (2019). Can job search assistance improve

the labour market integration of refugees? evidence from a field experiment. *Labour Economics 61*, 101745.

Behaghel, L., B. Crépon, and M. Gurgand (2014). Private and Public Provision of Counseling to Job Seekers: Evidence from a Large Controlled Experiment. *American Economic Journal: Applied Economics 6*(4), 142–174.

Bennmarker, H., E. Grönqvist, and B. Öckert (2013). Effects of contracting out employment services: Evidence from a randomized experiment. *Journal of Public Economics 98*, 68–84.

Bergman, M. A., P. Johansson, S. Lundberg, and G. Spagnolo (2016). Privatization and quality: Evidence from elderly care in sweden. *Journal of health economics 49*, 109–119.

Chan, D. C., D. Card, and L. Taylor (2023, November). Is there a va advantage? evidence from dually eligible veterans. *American Economic Review 113*(11), 3003–43.

Crépon, B. (2018). Private providers of labor market services: A review of the evidence.

Dahlberg, M., J. Egebark, U. Vikman, G. Özcan, et al. (2024). Labor market integration of refugees: RCT evidence from an early intervention program in Sweden. *Journal of Economic Behavior and Organization 217*, 614–630.

Duggan, M., J. Gruber, and B. Vabson (2018, February). The consequences of health care privatization: Evidence from medicare advantage exits. *American Economic Journal: Economic Policy 10*(1), 153–86.

Fabre, A. and S. Straub (2023). The Impact of Public–Private Partnerships (PPPs) in Infrastructure, Health, and Education. *Journal of Economic Literature 61*(2), 655–715.

Foged, M., L. Hasager, and G. Peri (2024). Comparing the effects of policies for the labor market integration of refugees. *Journal of Labor Economics 42*(S1), S335–S377.

Foged, M., L. Hasager, G. Peri, J. N. Arendt, and I. Bolvig (2022). Language Training and Refugees' Integration. *The Review of Economics and Statistics*, 1–41.

Foged, M. and C. van der Werf (2022). Access to language training and local integration of refugees. CReAM DP No. 10/22.

Fraser, A., S. Tan, M. Lagarde, and N. Mays (2018). Narratives of promise, narratives of caution: A review of the literature on social impact bonds. *Social policy & administration 52*(1), 4–28.

Gradus, R., M. Schoute, and E. Dijkgraaf (2018). The effects of market concentration on costs of local public services: empirical evidence from Dutch waste collection. *Local Government Studies 44*(1), 86–104.

Grossman, S. J. and O. D. Hart (1983). An analysis of the principal-agent problem. *Econometrica 51*(1), 7–45.

Hart, O., A. Shleifer, and R. W. Vishny (1997). The Proper Scope of Government: Theory and an Application to Prisons. *The Quarterly Journal of Economics 112*(4), 1127–1161.

Heller, B. H. and K. S. Mumma (2023). Immigrant integration in the united states: The role of adult english language training. *American Economic Journal: Economic Policy 15*(3), 407–37.

Holmström, B. (1979). Moral hazard and observability. *The Bell Journal of Economics*, 74–91.

Holmström, B. and P. Milgrom (1991). Multitask principal–agent analyses: Incentive contracts, asset ownership, and job design. *The Journal of Law, Economics, and Organization 7*(special_issue), 24–52.

Joona, P. A. and L. Nekby (2012). Intensive Coaching of New Immigrants: An Evaluation Based on Random Program Assignment*. *The Scandinavian Journal of Economics 114*(2), 575–600.

Karinen, R., N. Koivula, T. Koivula, O. Oosi, H. Pesola, M. Sarvimäki, and T. Virkola (2024). Impacts and practical implementation of the koto-sib experiment. Publications of the Ministry of Economic Affairs and Employment, 2024:34.

Knutsson, D. and B. Tyrefors (2023). The Quality and Efficiency of Public and Private Firms: Evidence from Ambulance Services. *The Quarterly Journal of Economics 137*(4), 2213–2262.

Krug, G. and G. Stephan (2016). Private and public placement services for hard-to-place unemployed: Results from a randomized field experiment. *ILR Review 69*(2), 471–500.

Lochmann, A., H. Rapoport, and B. Speciale (2019). The effect of language training on immigrants' economic integration: Empirical evidence from france. *European Economic Review 113*, 265–296.

OECD (2016). Social Impact Bonds: State of Play and Lessons Learnt. OECD Working Paper.

Pesola, H. and M. Sarvimäki (2024). Intergenerational Spillovers of Integration Policies: Evidence from Finland's Integration Plans. *Review of Economics and Statistics, forthcoming*.

Pesola, H., M. Sarvimäki, and T. Virkola (2024). Paths to Integration: Labor Market and Social Integration Across Immigrant Admission Categories.

Rehwald, K., M. Rosholm, and M. Svarer (2017). Do public or private providers of employment services matter for employment? Evidence from a randomized experiment. *Labour Economics 45*, 169–187.

Romero, M., J. Sandefur, and W. A. Sandholtz (2020). Outsourcing education: Experimental evidence from Liberia. *American Economic Review 110*(2), 364–400.

Sarvimäki, M. and K. Hämäläinen (2016). Integrating Immigrants: The Impact of Restructuring Active Labor Market Programs. *Journal of Labor Economics 34*(2), 479–508.

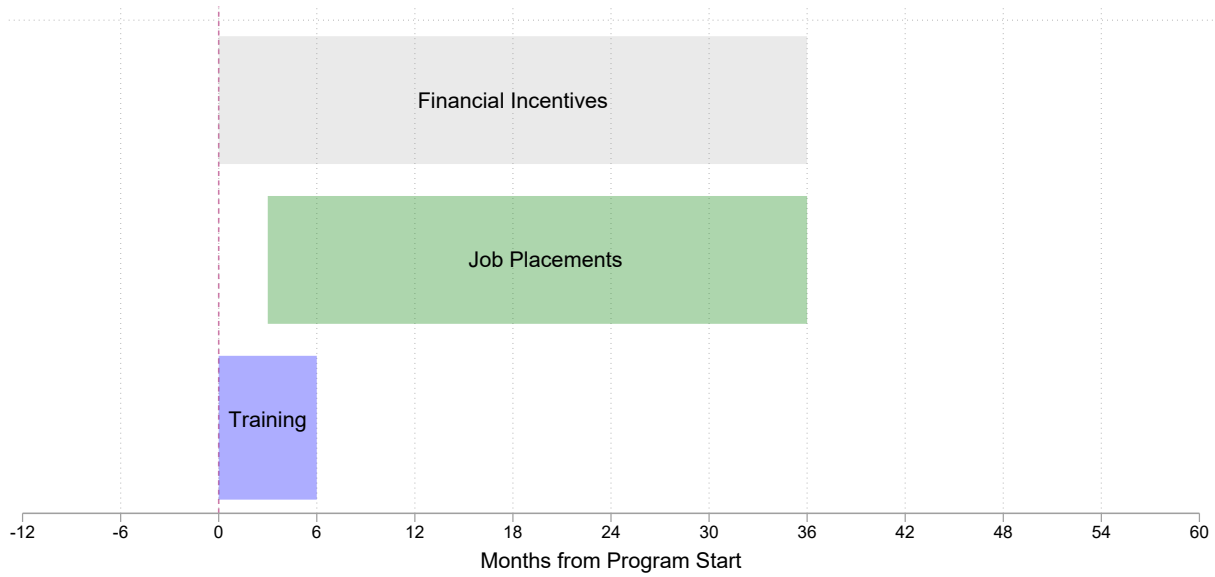
Shleifer, A. (1985). A theory of yardstick competition. *The RAND journal of Economics*, 319–327.

Szymanski, S. (1996). The impact of compulsory competitive tendering on refuse collection services. *Fiscal Studies 17*(3), 1–19.

Szymanski, S. and S. Wilkins (1993). Cheap rubbish? competitive tendering and contracting out in refuse collection–1981–88. *Fiscal Studies 14*(3), 109–130.

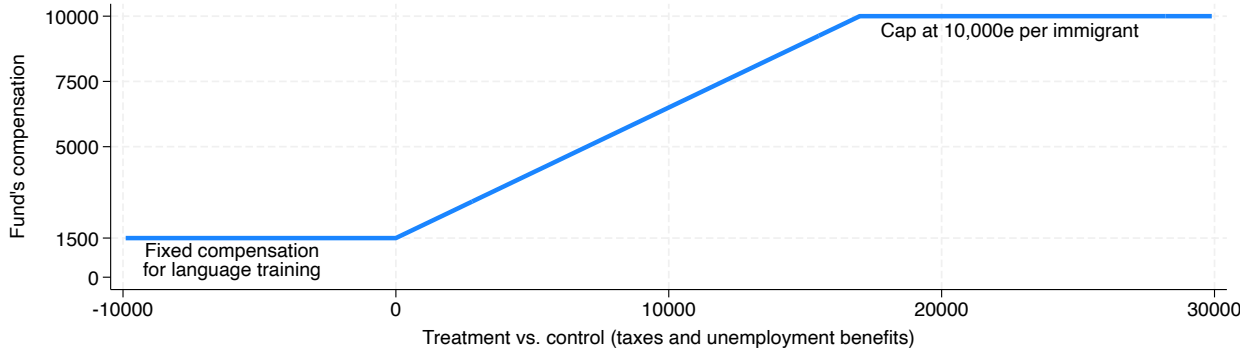
Figures and Tables

Figure 1: Timeline

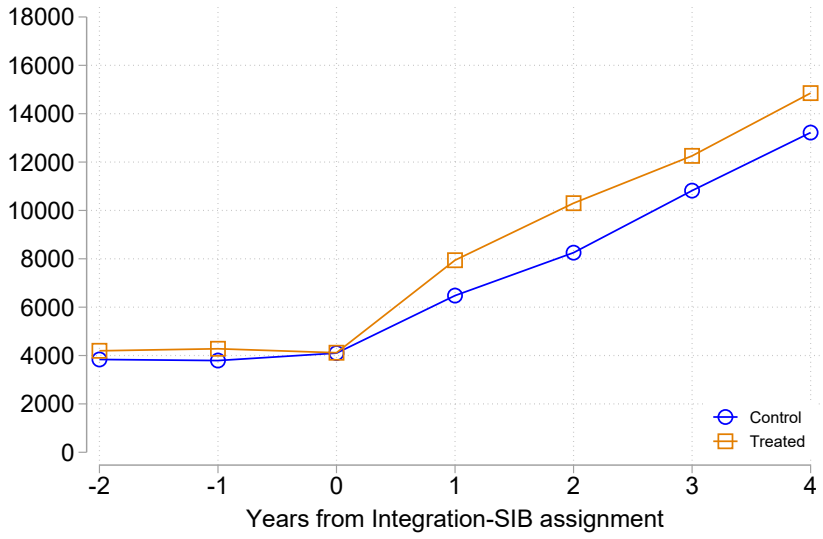


Notes: This figure shows the timeline of the Integration SIB for the participants. After being selected into the treatment group, participants are subject to a 7-week intensive training that aims to provide practical language skills suitable for the target job. After this trainign program the private provider aims to match the participants to a job. The service provider has financial incentives to maximize savings during the 36 months after assignment.

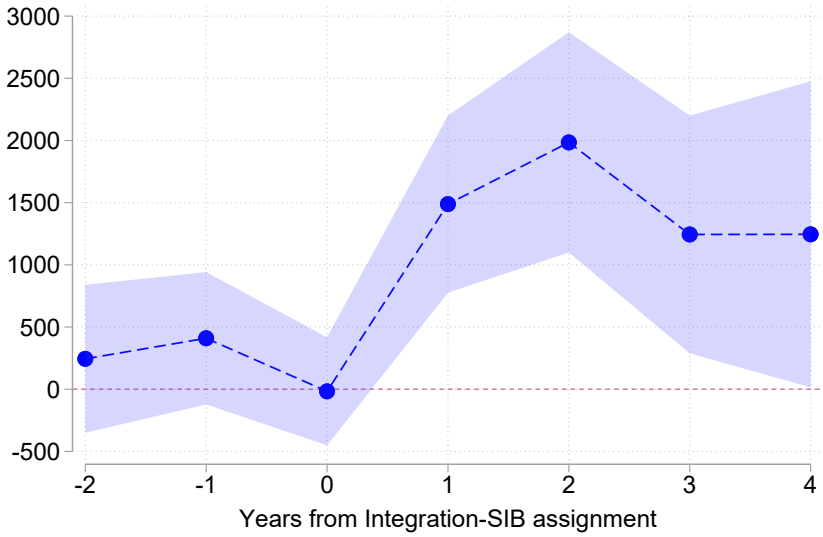
Figure 2: Financial Incentives for the Private Fund



Notes: This figure shows the financial incentives for the private. Their compensation is increasing in the difference.



(a) Trend: Annual Labor Earnings

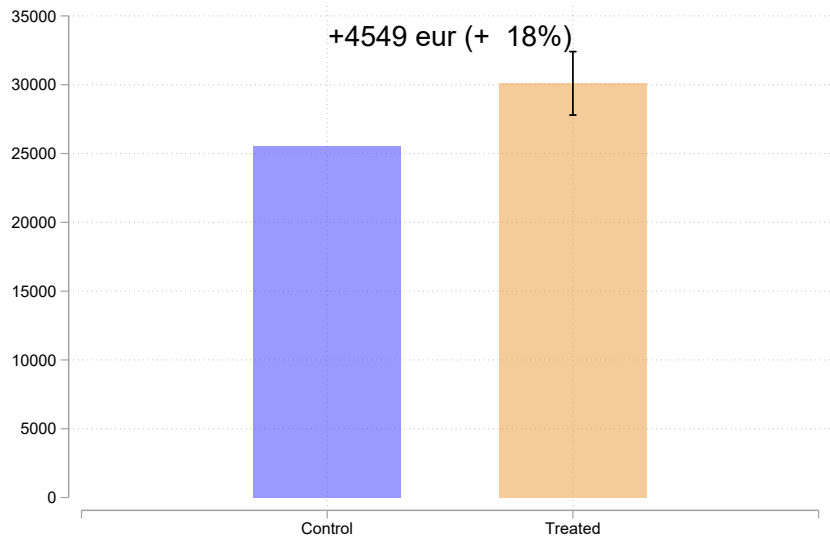


(b) Difference: Annual Labor Earnings

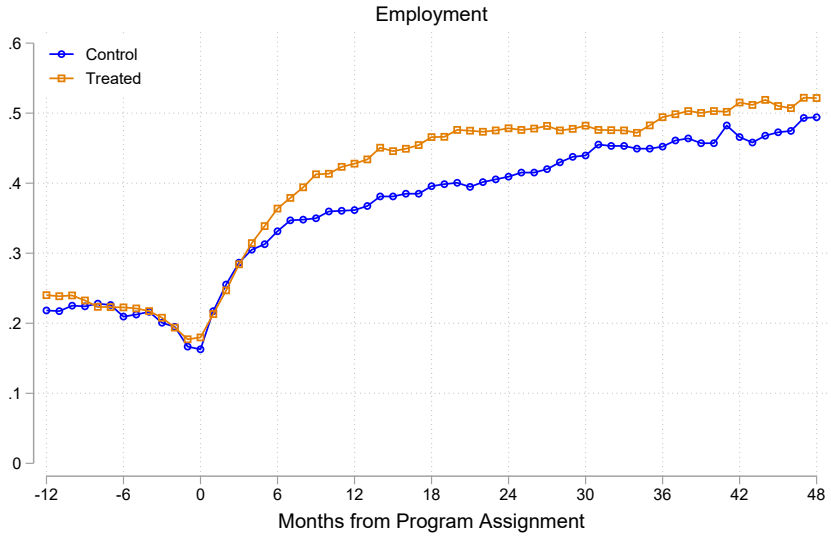
Figure 3: The Effects of the Integration SIB on Earnings

Note: This figure shows differences in earnings among the treatment group and control groups in the Integration Sib -program. Panel (a) shows the trends in annual earnings in the years around the program assignment. Panel (b) shows the treatment effects for each year separately.

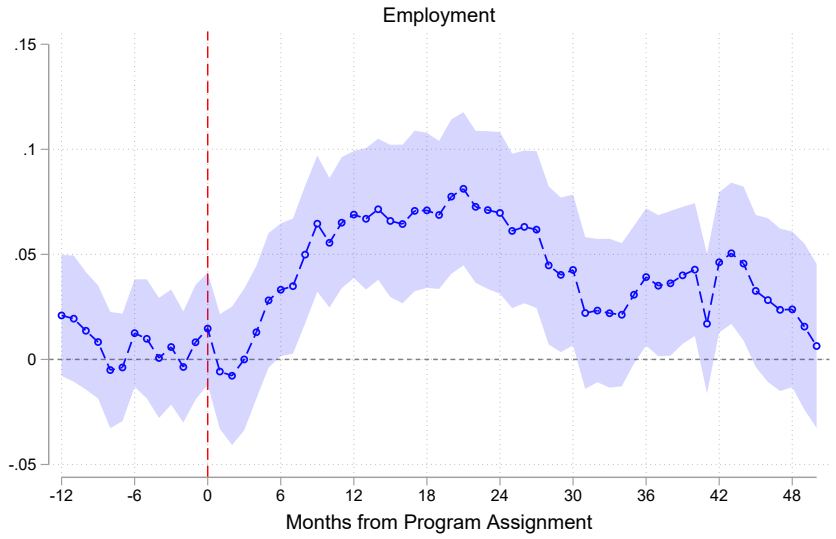
Figure 4: The Effects of the Integration SIB on Total Earnings



Notes: This figure shows the cumulative earnings during the three years following the program start by treatment status of participants. Black lines show the confidence intervals for the point estimates of the difference between treatment and control groups..



(a) Trend: Monthly Employment

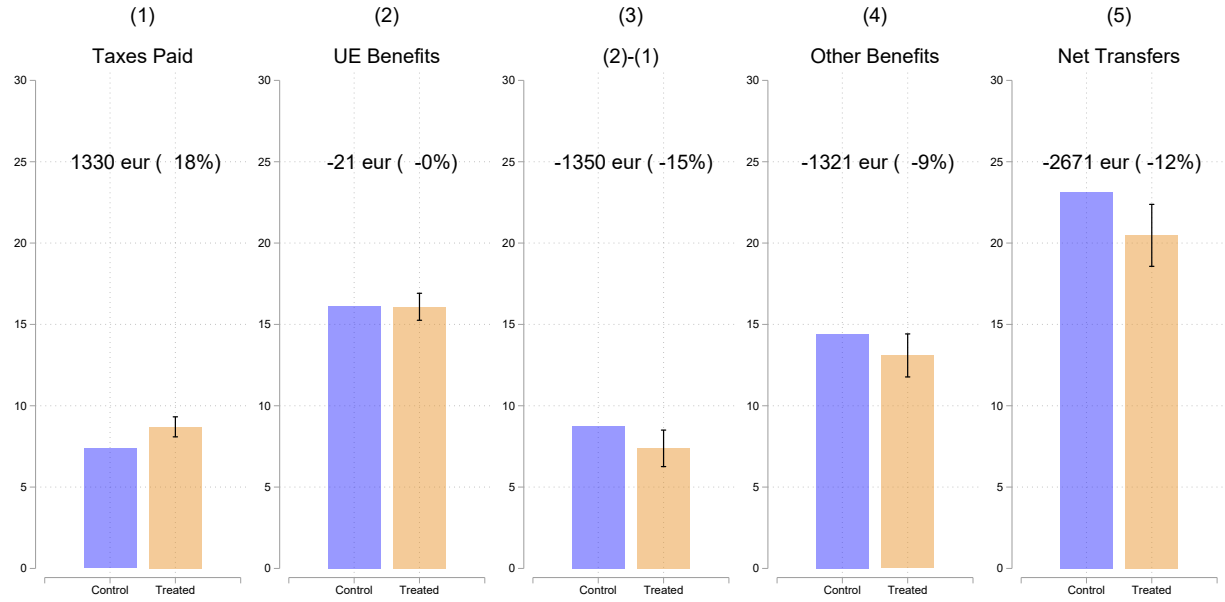


(b) Difference: Monthly Employment

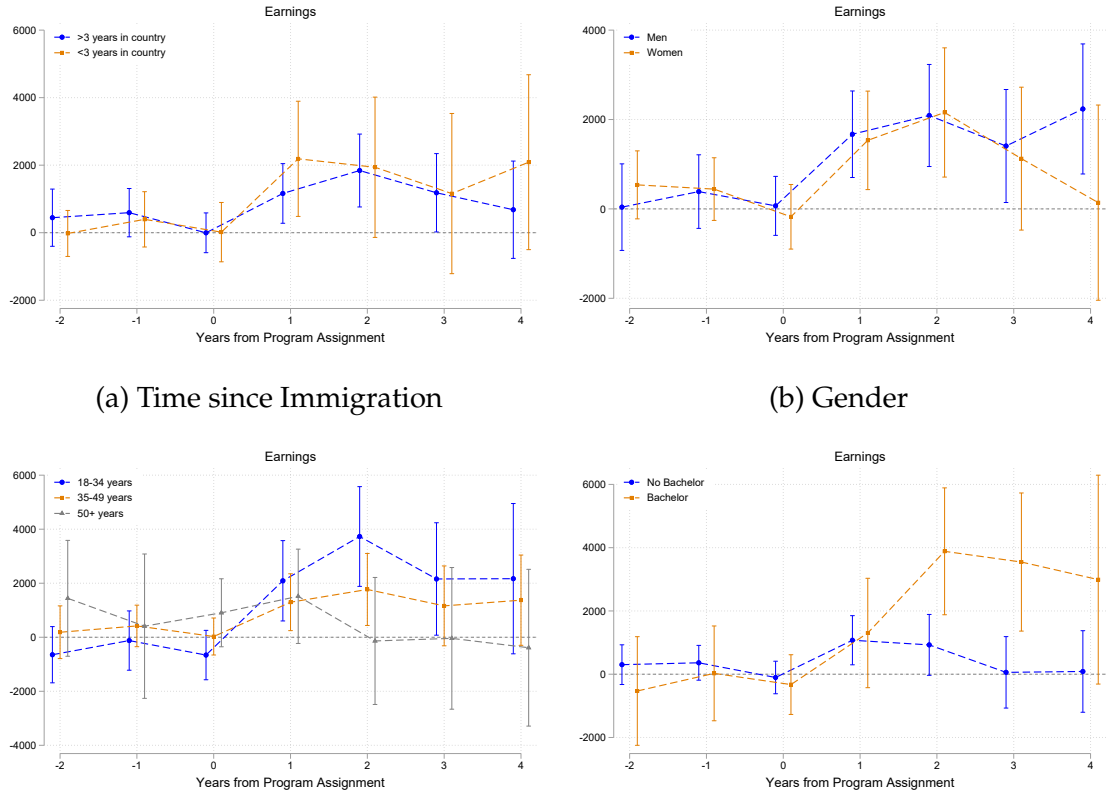
Figure 5: The Effects of the Integration SIB on Employment

Note: This figure shows differences in employment among the treatment group and control groups in the Integration SIB -program. Panel (a) shows the trends in employment in the months around the program assignment. Panel (b) shows the monthly treatment effects

Figure 6: Decomposition of Net Transfers among Integration SIB participants



Notes: This figure shows the decomposition of net transfers to Integration SIB-participants during the three years following the program start. Blue bars show the outcomes for the control group and yellow bars show the outcomes for the treatment group. The first column from the left hand side shows taxes and social security contributions paid by individuals in the two groups. Black lines show the 95 percent confidence intervals for the difference in point estimates. Second column shows unemployment benefits collected. Third column shows the difference between the two, i.e. net unemployment benefits. Fourth column shows other social security benefits collected that were not part of the private providers incentives. Fifth column shows the total net transfers including all social benefits minus taxes and social security contributions.



(a) Time since Immigration

(b) Gender

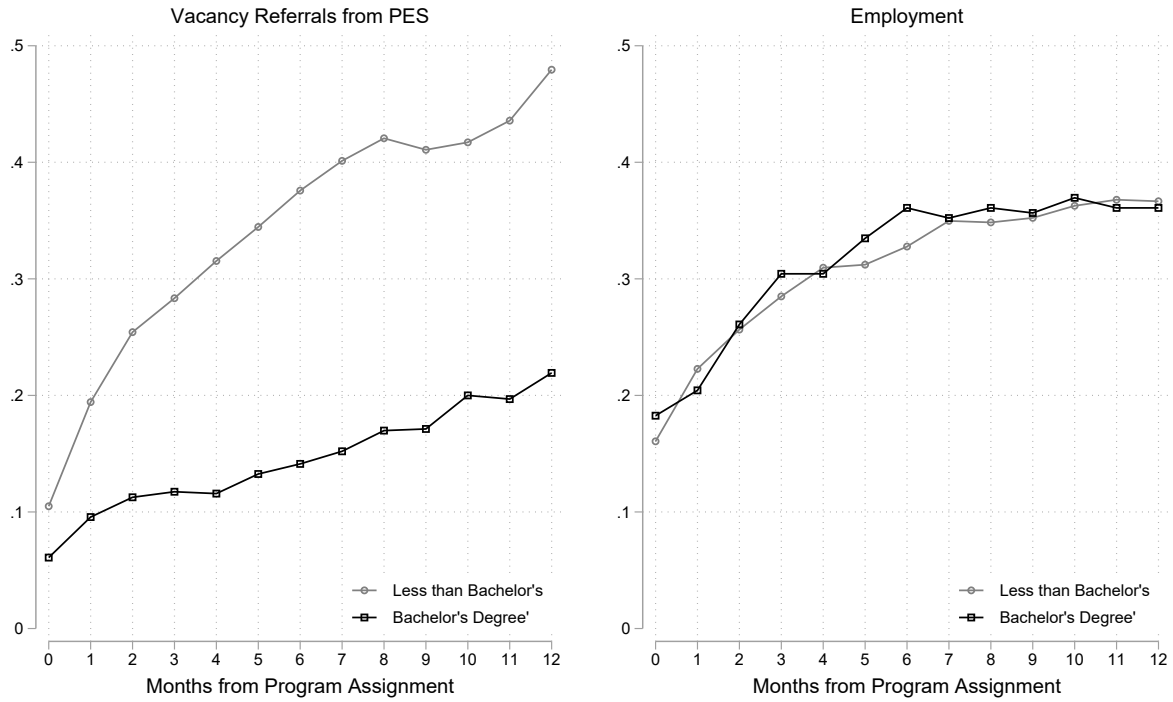
(c) Age

(d) Education

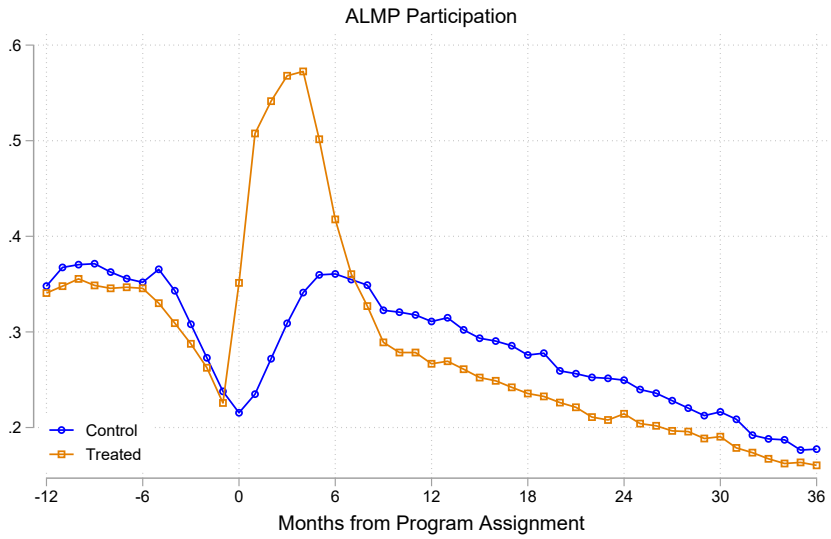
Figure 7: Event-Study Estimates by Participant Characteristics

Note: This figure shows the differences in earnings between the treatment group and the control group as a function of years from program assignment. Each estimate comes from a separate regression with randomization cluster fixed effects and standard errors are clustered at randomization event level.

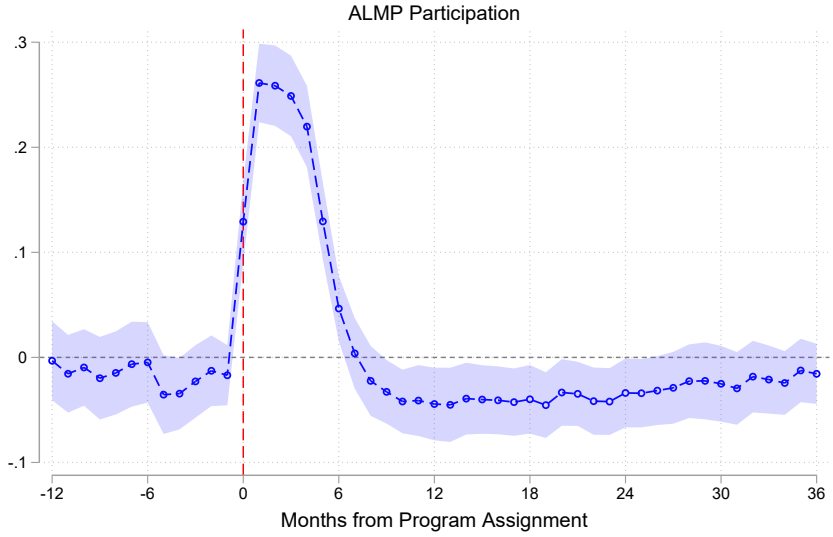
Figure 8: Vacancy Referrals and Employment in the Control Group



Notes: This figure shows the incidence of vacancy referrals and employment trajectories by the PES to job seekers in the control group by skill groups. The left hand side shows the cumulative number of vacancy referrals per job seeker in the 12 months following the program assginment. The right hand side shows the employment rate.



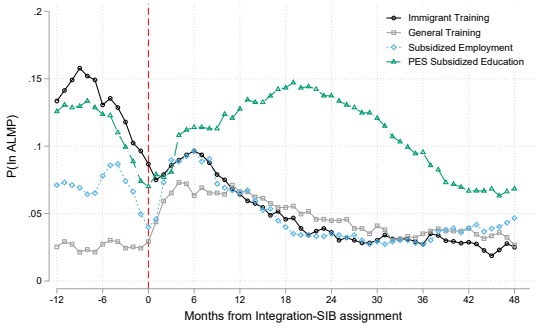
(a) Trend: P(Any ALMP)



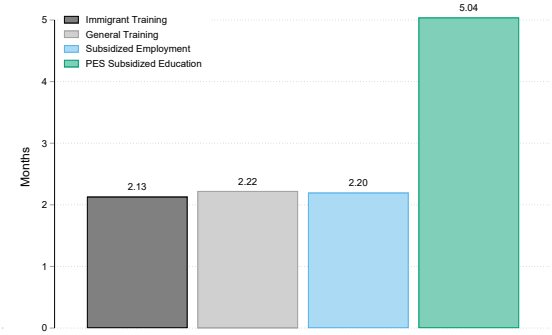
(b) Difference: P(Any ALMP)

Figure 9: Participation in Active Labor Market Programs

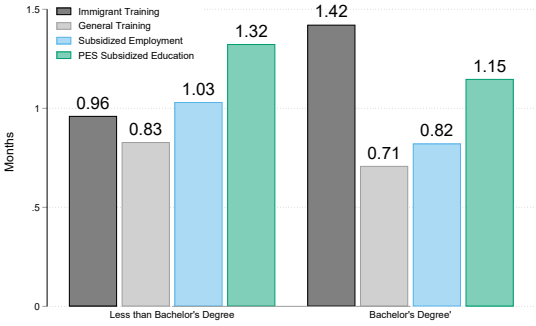
Note: This figure shows trends for treatment and control groups (panel a) and treatment effects (panel b) in active labor market program participation.



(a) Over Time



(b) Cumulative Months After Assignment

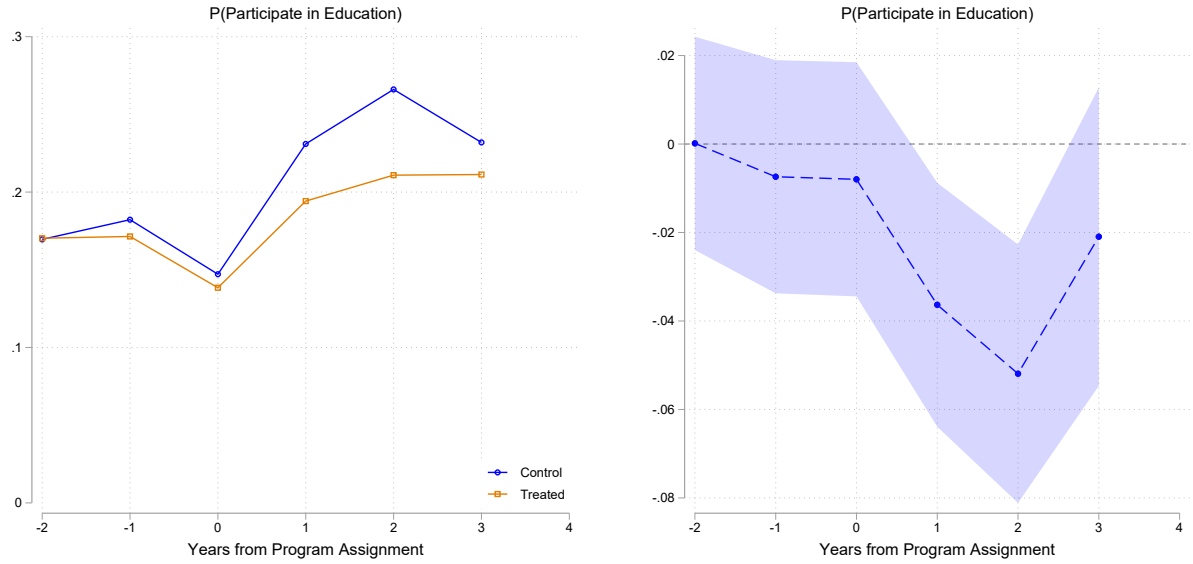


(c) PES Supported Training (0-12 Months)

Figure 10: PES Services for Control Group

Note: This figure shows participation in active labor market programs for the control group. Panel (a) shows the shares participating in different services in the months around the program assignment. Panel (b) shows the cumulative months in each type of service following program assignment

Figure 11: Participation in Education Programs



Notes: This figure shows the differences in enrollment in general education. Majority of these consist of vocational secondary education

Table 1: Pre-Treatement Balance

	(1) Control	(2) Treated	(3) β^{SIB}	(4) SE
Assignment Year	2018.28	2018.30	-0.00	(0.00)
Age	38.50	38.85	0.43	(0.35)
Woman	0.41	0.42	0.00	(0.02)
Married	0.56	0.59	0.02	(0.02)
Single	0.25	0.22	-0.03**	(0.01)
Divorced	0.17	0.18	0.01	(0.01)
Years in Country	6.87	6.84	0.08	(0.19)
Days Unemployed	214	232	18*	(11)
Earnings (t-1)	3792	4279	446	(297)
Social Benefits (t-1)	10394	9990	-274	(293)
Unemployment Benefits (t-1)	5749	5639	-46	(166)
Net Transfers (t-1)	-8759	-8286	346	(300)
Work Days (t-1)	74.82	82.54	7.43*	(4.49)
Enrolled in Education Program (t-1)	0.18	0.17	-0.01	(0.01)
Enrolled in Secondary Program (t-1)	0.15	0.14	-0.01	(0.01)
N	1026	2636		

Note: This table describes the differences between the treatment group and the control group at the time of randomization. Column (1) shows the characteristics fo the control group and Column (2) the characteristics of the treatment group. Column (3) shows the differences between the treatment and control group, conditional on randomization cluster fixed effects and column (4) shows the standard errors with randomization event clustering.

Table 2: Treatment Effects by Intensity of Program Roll-Out

			By Region	
	(1) Pooled	(2) Pooled	(3) Uusimaa	(4) Rest-of-Finland
Panel A: Earnings				
Treated	6489.3**	5355.0**	4615.2***	5355.0**
	(3089.3)	(2104.8)	(1321.3)	(2162.9)
Treated X Intensity	-559.6			
	(959.8)			
Treated X Uusimaa		-739.9		
		(2487.2)		
Outcome mean	29180	29193	29583	27286
N	3,645	3,640	3,022	618
Panel B: Months of Employment				
Treated	2.578*	2.437**	1.640***	2.437**
	(1.416)	(1.062)	(0.500)	(1.091)
Treated X Intensity	-0.276			
	(0.425)			
Treated X Uusimaa		-0.798		
		(1.174)		
Outcome mean	15.063	15.069	15.107	14.880
N	3,645	3,640	3,022	618

Note: This table shows how the labor market outcomes of participants vary across the intensity of the program roll-out in the labor market. In column (1), we interact the treatment status with the share of labor force in the labor market being part of the program. In column (1), we interact treatment status with a dummy on being in the Uusimaa reion where the roll out was most significant and where more than 80 percent of the participants fell in. In columns (3) and (4) we show the effects separately for Uusimaa region and the rest of Finland. We control for participant age and gender in all specifications.

Table 3: The Effects on Types of Jobs

	Share working in occupation			Skill content	
	(1) Managers and professionals	(2) Clerical and service	(3) Manual and elementary	(4) Occ. Earnings	(5) Occ. High Edu
Panel A: Average Treatment Effects					
Treated	0.0228***	0.00730	0.00544	1229.1***	0.0281***
	(0.00759)	(0.00750)	(0.0124)	(423.3)	(0.00885)
Panel B: Treatment Effects by Job Seeker's Skill					
Treated	0.004	0.006	0.011	150	0.005
	(0.006)	(0.009)	(0.015)	(380)	(0.008)
Treated X HighEdu	0.055**	-0.010	-0.008	2726**	0.054*
	(0.026)	(0.022)	(0.027)	(1264)	(0.030)
Mean	0.074	0.106	0.216	29304	0.159
N	10952	10952	10952	4071	4071

Note: This table shows the treatment effects on occupations. Columns (1)-(3) shows the effects on the probability of working in different types of occupations and Columns (4) and (5) shows the effects on measures of occupational skill content. Occupation earnings measures the average earnings of all Finnish workers in each occupation. Occupation High Edu measures the share of all Finnish workers in that occupation who hold at least a Bachelor's degree. We control for randomization cluster fixed effects and participant age and gender. Standard errors are clustered at randomization cluster level.

Table 4: Treatment Effects by Participant Characteristics

	(1)	(2)	(3)	(4)
Panel A: Earnings				
Treated	4367***	5036***	4694***	2243*
	(1405)	(1541)	(1168)	(1239)
Treated X Recent	1173			
	(3246)			
Treated X Woman		-772		
		(2386)		
Treated X Age			-226*	
			(136)	
Treated X High Edu				7893***
				(2709)
Mean	28936	29181	29181	29177
N	3426	3645	3645	3550
Cluster FE	✓	✓	✓	✓
Panel B: Employment (days)				
Treated	56***	52***	57***	48***
	(15)	(17)	(12)	(13)
Treated X Recent	-7			
	(31)			
Treated X Woman		13		
		(27)		
Treated X Age			-2	
			(1)	
Treated X High Edu				21
				(27)
Mean	386	393	393	391
N	3426	3645	3645	3550
Cluster FE	✓	✓	✓	✓

Note: This table shows how the labor market outcomes of participants vary across by pre-registered participant characteristics. In column (1), we interact the treatment status with recent immigrant status, i.e., the participant has been in the country less than three years at the time of program assignment and is thus eligible for government integration services. Columns (2)-(4) show similar estimates where we interact treatment status with participants' gender, age and higher education status. We control for participant age and gender in all specifications.

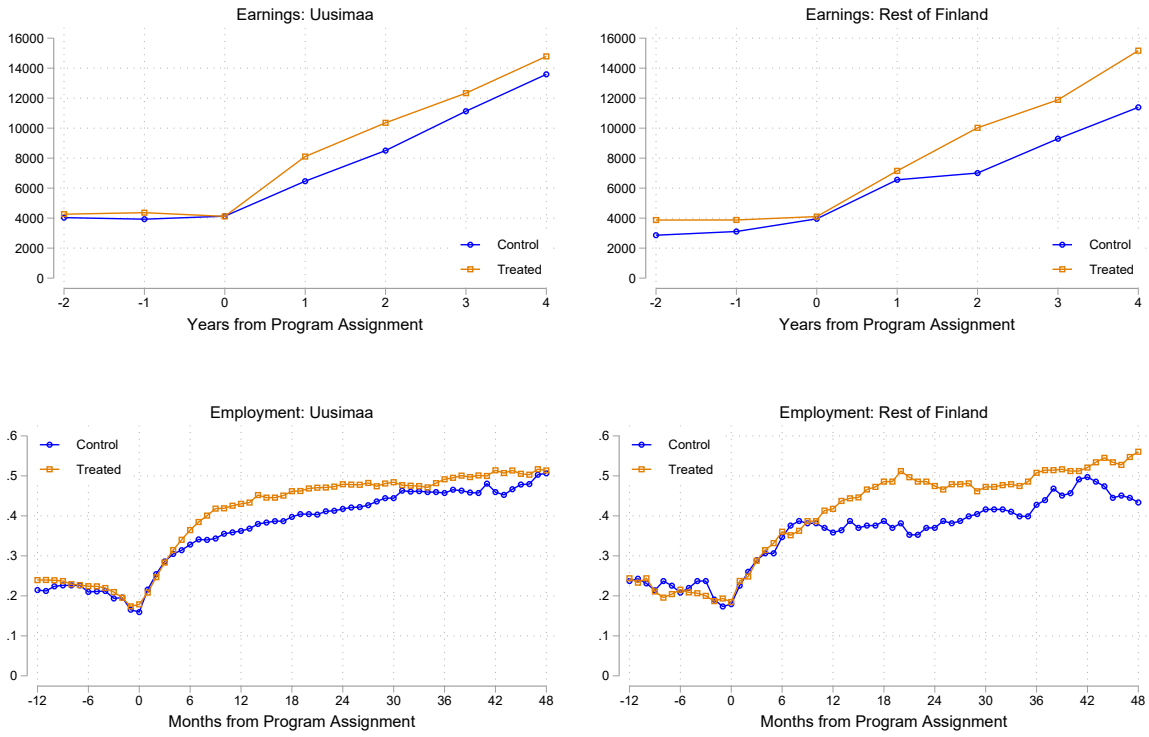
Table 5: PES Makes Vacancy Referrals to Low Ranked Occupations

	(1) Managers and Professionals	(2) Clerical and Service	(3) Manual and Elementary	(4) Total
Panel A: All participants				
Jobs (Treated)	0,08	0,11	0,22	0,41
Jobs (Control)	0,05	0,10	0,22	0,37
Share of Jobs (Control)	0,14	0,27	0,58	1,00
No. Referrals	0,02	0,08	0,18	0,28
Share of Referrals	0,05	0,30	0,65	1,00
Panel B: Low-skilled				
Jobs (Treated)	0,03	0,11	0,25	0,39
Jobs (Control)	0,03	0,11	0,24	0,38
Share of jobs (Control)	0,08	0,28	0,64	1,00
No. Referrals (Control)	0,01	0,10	0,22	0,34
Share of Referrals (Control)	0,03	0,31	0,66	1,00
Panel C: High-skilled				
Jobs (Treated)	0,20	0,10	0,14	0,44
Jobs (Control)	0,13	0,11	0,13	0,37
Share of jobs (Control)	0,35	0,30	0,35	1,00
No. Referrals (Control)	0,03	0,02	0,05	0,09
Share of Referrals (Control)	0,29	0,19	0,52	1,00

Note: This table shows jobs and PES vacancy referrals to program participants by occupations. We show these separately for all participants and by skill of the participant. Column (1) shows referrals to professional and managerial occupations, Column (2) to Clerical and Service jobs and (3) to manual an elementary occupations. Column (4) shows total jobs and referrals.

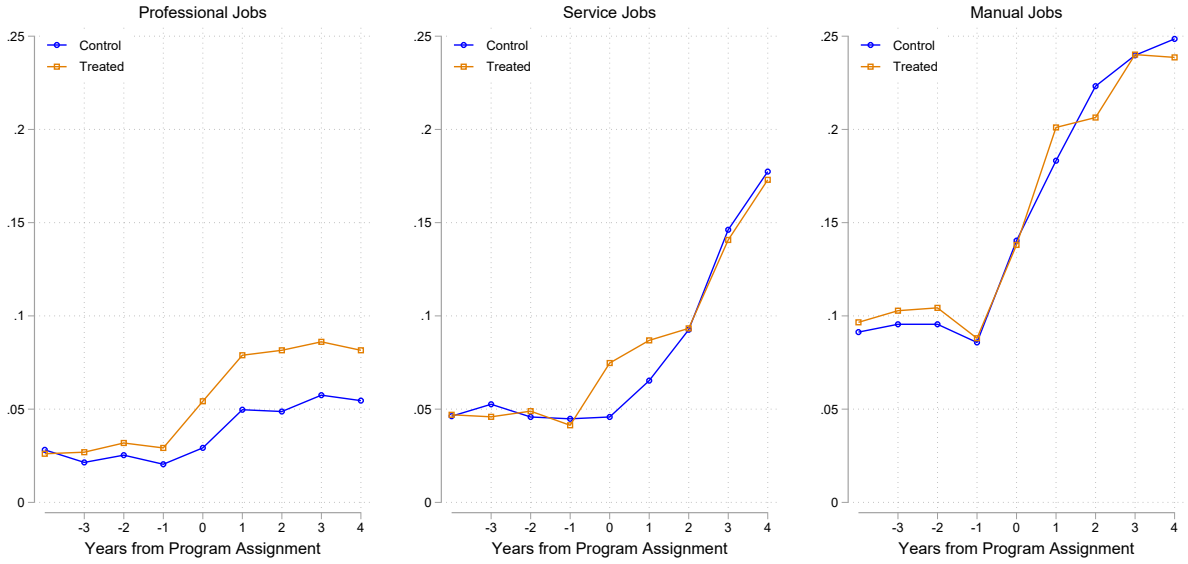
Appendix

Figure A1: Lack of Displacement: Weakly Decreasing in Treatment Intensity



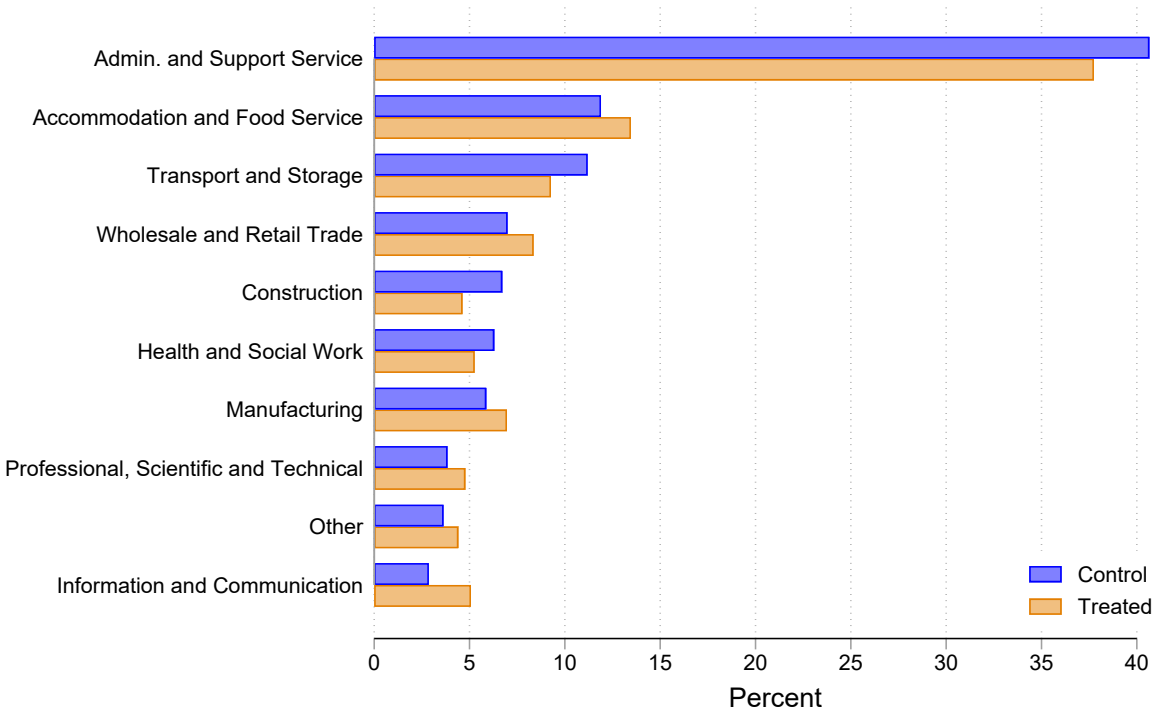
Notes: This figure shows trends in earnings and employment by the labor markets with differences in the intensity of the program roll-out.

Figure A2: Occupations Held by Program Participants



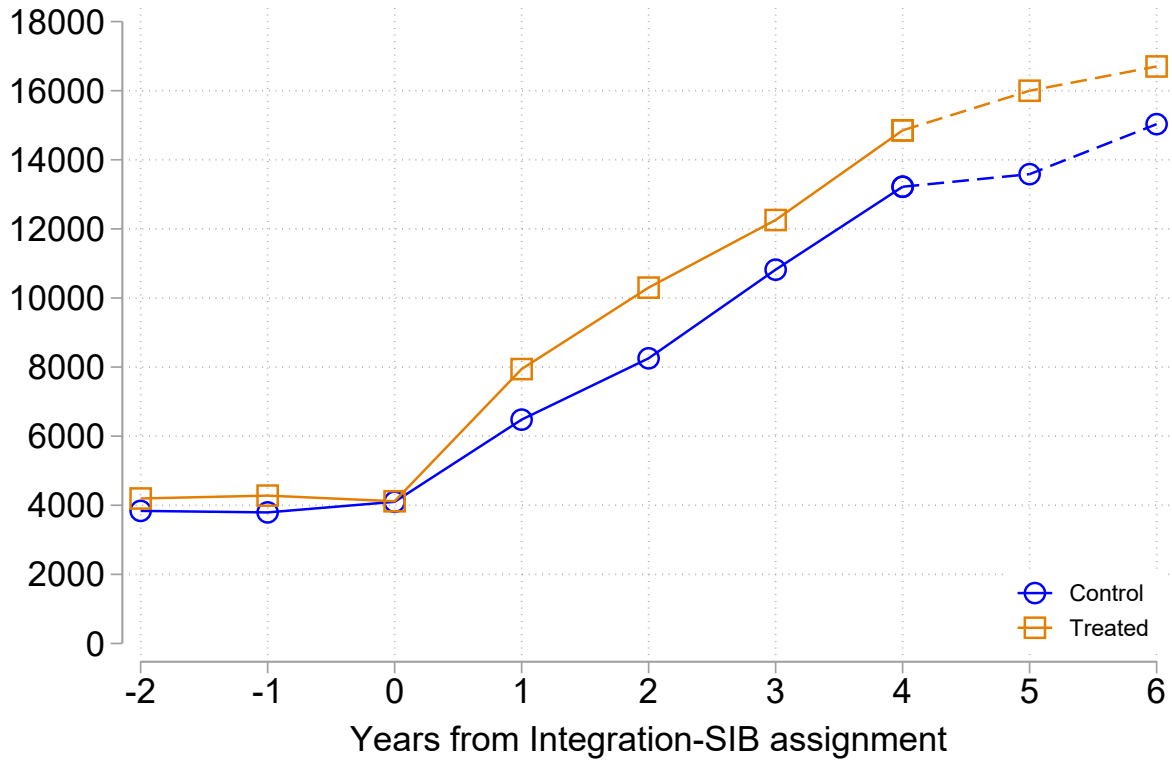
Notes: This figure shows the differences in occupations help by control and treatment groups around program assignment. We classify jobs based on 1-digit ISCO codes. Professional jobs include groups 1-3, including both managers and professional jobs. Service jobs include service sector and clerical jobs (4-5) and Manual jobs include groups 6-9.

Figure A3: Industry Distribution Shifts to ICT/Manufacturing/Restaurants/Retail



Notes: This figure shows the distribution of jobs in the treatment and control group across broad industries.

Figure A4: Long-Run Earnings

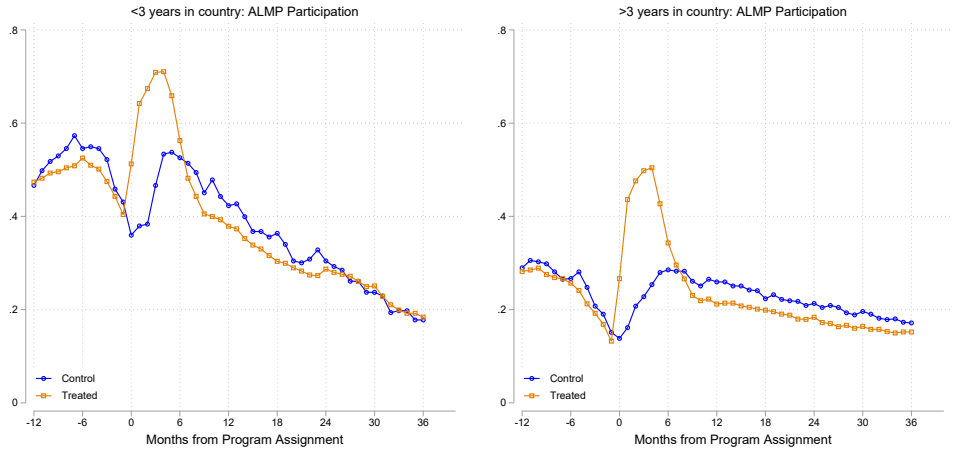


Notes: This figure shows earnings for treatment and control groups. For years 5 and 6 after program assginment the sample consists of the first two and the first cohort only respectively to allow for longer follow-up period.

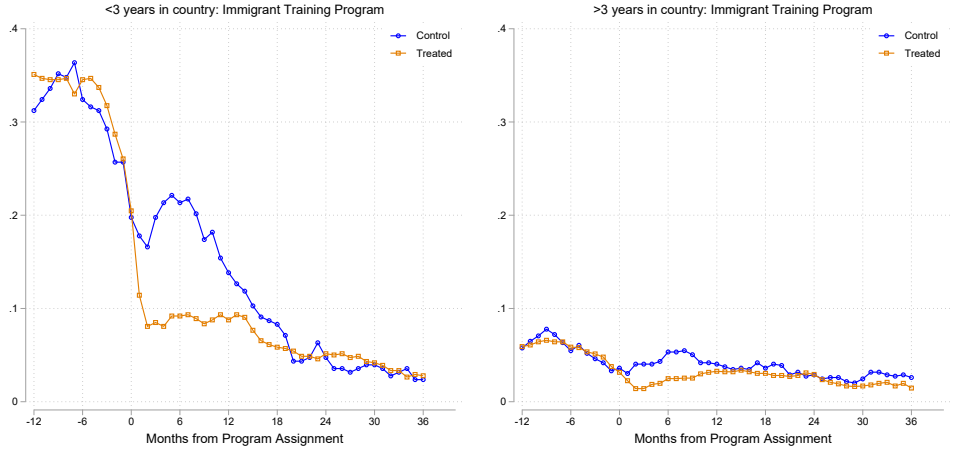
Table A1: Program Roll-Out Across Labor Markets

Region	(1) Labor Force	(2) Immi-grants	(3) Immi-grant Share	(4) Partici-pants	(5) Partici-pants per 1000	(6) Partic-pants per 1000 Immi-grants
Uusimaa	843571	64704	0.08	3034	3.597	46.890
Varsinais-Suomi	227000	9362	0.04	253	1.115	27.024
Pohjois-Karjala	74397	1606	0.02	73	0.981	45.455
Pirkanmaa	245371	7058	0.03	187	0.762	26.495
Pohjois-Pohjanmaa	186439	3388	0.02	64	0.343	18.890
Pohjanmaa	115360	4969	0.04	23	0.199	4.629
Kaakkois-Suomi	136679	5509	0.04	17	0.124	3.086
Keski-Suomi	126637	2525	0.02	11	0.087	4.356
Satakunta	101175	2677	0.03	0	0.000	0.000
Häme	178050	5365	0.03	0	0.000	0.000
Etelä-Savo	61888	1293	0.02	0	0.000	0.000
Pohjois-Savo	114775	2379	0.02	0	0.000	0.000
Etelä-Pohjanmaa	89165	1770	0.02	0	0.000	0.000
Kainuu	33192	595	0.02	0	0.000	0.000
Lappi	82528	1698	0.02	0	0.000	0.000
Ahvenanmaa	15094	1711	0.11	0	0.000	0.000

Note: This table shows the share of participants in the Integration SIB program by Public Employment Service regions. This variation is driven by selective roll-out by Public Employment Services



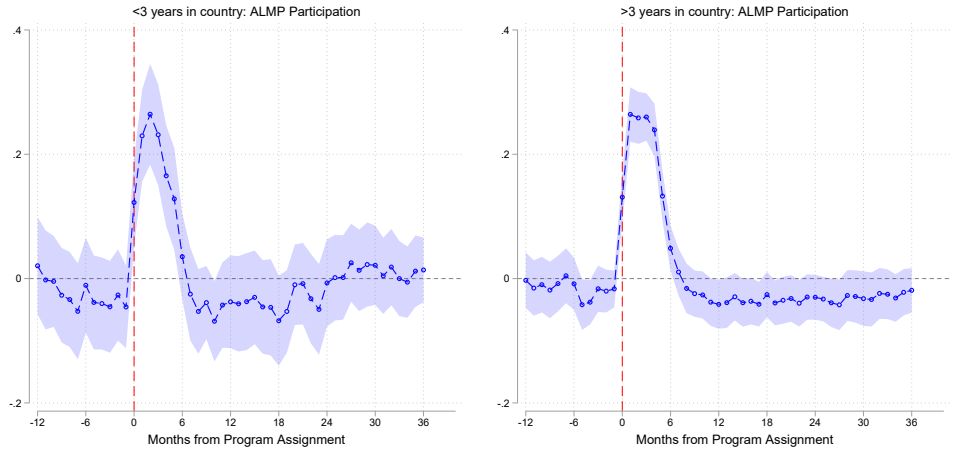
(a) Trend: P(Any ALMP)



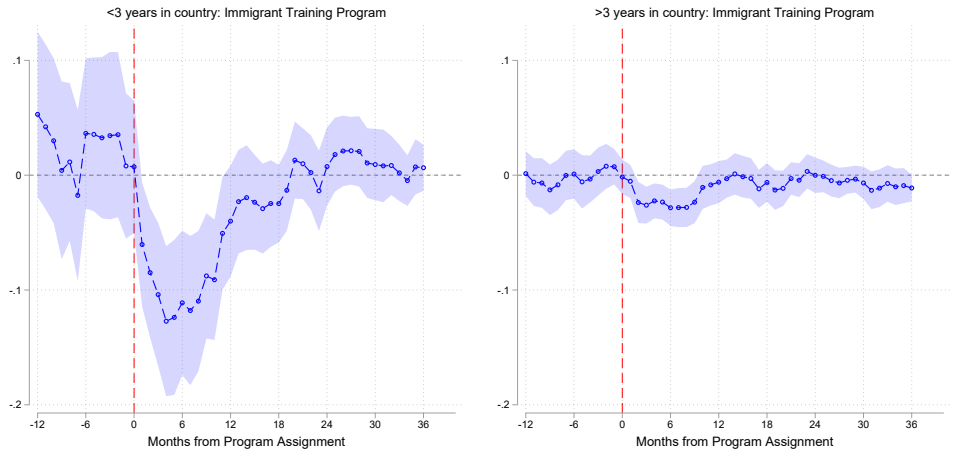
(b) Trend: P(Immigrant Training)

Figure A5: Active Labor Market Programs by Years in Country

Note: This figure shows trends in active labor market program participation separately for participants that have been either less than or more than three years in the country and are thus differentially subject to integration training plans. Panel (a) shows overall participation in any active labor market programs. Panels (b) shows participation in language and integration training targeted to unemployed immigrant job seekers.



(a) Difference: P(Any ALMP)



(b) Difference: P(Immigrant Training)

Figure A6: Active Labor Market Programs by Years in Country

Note: This figure shows the treatment effects in active labor market program participation separately for participants that have been either less than or more than three years in the country and are thus differentially subject to integration training plans. Panel (a) shows overall participation in any active labor market programs. Panels (b) shows differences in language and integration training targeted to unemployed immigrant job seekers.

Table A2: The Effects on Co-worker and Firm Characteristics

	Co-worker Earnings		AKM FE's		Firm Productivity	
	(1) Level	(2) Log	(3) Firm	(4) Co-worker	(5) log(Sales per Worker)	(6)
Panel A: Average Treatment Effects						
Treated	1511.7***	0.0531*	0.000912	0.00234	0.0877**	0.0553*
	(579.6)	(0.0284)	(0.00759)	(0.00801)	(0.0382)	(0.0288)
Mean	22506	9.799	-0.074	-0.226	11.334	11.334
N	6599	6599	5771	6340	5407	5407
Panel B: Effects by Skill						
Treated	68.8	0.024	-0.009	-0.013	0.057	0.034
	(636.6)	(0.036)	(0.009)	(0.009)	(0.047)	(0.035)
Treated X HighEdu	4857.0***	0.090	0.026	0.054***	0.120	0.095
	(1430.1)	(0.063)	(0.017)	(0.019)	(0.089)	(0.069)
Mean	22506	9.799	-0.074	-0.226	11.334	11.334
N	6409	6409	5603	6159	5256	5256
Industry FE						Yes

Note: This table shows the treatment effects on workplace and firm characteristics. Column (1) shows the effect on co-worker earnings and column (2) on log earnings. Column (3) shows the effect on firm FE estimated from an AKM model on earnings with person and firm fixed effects between 2014–2019. Column (4) shows the estimated co-worker FE's from the same model. Columns (5) and (6) show the effects on firm productivity as measured by log sales per worker. We control for randomization cluster FE and participant age and gender in all specifications. Standard errors are clustered at the randomization event level.

Table A3: The Effects on Workplace Segregation

	Co-Workers		Managers	
	(1) Immigrants	(2) Same-origin	(3) Immigrants	(4) Same-origin
Panel A: Average Treatment Effects				
Treated	-0.00286	0.00294	-0.00707	0.000405
	(0.0108)	(0.00761)	(0.0140)	(0.0100)
Outcome mean	0.334	0.086	0.221	0.097
N	6599	6599	6706	6706
Panel B: Treatment Effects by Skill				
Treated	0.004	0.007	0.005	0.002
	(0.012)	(0.010)	(0.017)	(0.013)
Treated X HighEdu	-0.022	-0.013	-0.047	-0.008
	(0.024)	(0.020)	(0.036)	(0.024)
Mean	0.334	0.086	0.221	0.097
N	6409	6409	6514	6514

Note: This table shows the treatment effects on workplace segration. Columns (1) to (2) show the effect on co-workers' immigration status and origin country. Columns (3) and (4) show similar effects for manager characteristics. Managers are defined as the highest-paid workers in every firm. We control for randomization cluster FE and participant age and gender in all specifications. Standard errors are clustered at the randomization event level.

Table A4: Treatment Effects by Participant Citizenship

	(1) Earnings	(2) Net Transfers	(3) Employment
Treated	5711***	2532*	58***
	(1828)	(1351)	(22)
Treated X Russia	-7557*	-4844	-55
	(4422)	(3210)	(49)
Treated X Iraq	-2521	1688	30
	(4297)	(3792)	(52)
Treated X Somalia	-2276	-941	-16
	(3949)	(4076)	(54)
Treated X Estonia	4533	5798	73
	(4885)	(4005)	(55)
Treated X India	7537	-239	-3
	(8500)	(4930)	(70)
Treated X Thailand	-7194	-5783	34
	(5622)	(3926)	(69)
Treated X Turkey	5285	7145	-67
	(5875)	(4468)	(80)
Treated X China	-12470	-1506	-76
	(9505)	(5769)	(100)
Treated X Afghanistan	-13424	-8762	-32
	(8581)	(6286)	(79)
Treated X Iran	6336	14903***	86
	(6568)	(5364)	(79)
Mean	29352	-20990	395
N	3574	3574	3574
Cluster FE	✓	✓	✓

Note: This table shows the treatment effects by participant citizenship. We interact treatment status with the ten most common nationalities, which account for 47 percent of all participants. The immigrants from other sending countries are the reference group. Column (1) shows the effects on earnings, column (2) on net transfers and Column (3) on days of employment.